

# Control Engineering by Derivation

The state as a posterior, and the controller as a decision over it

Draft · 343 pages



**Ricardo Marquez, Ph.D.**

ricardo-marquez.com

# Control Engineering by Derivation

Volume I

From the control engineer's questions to probabilistic inference and decisions

Ricardo Marquez

November 2, 2026 (work in progress: chapters are added as they are written)

# How to read this book

This book is written for the engineer who designs, tunes and commissions control loops and the supervision above them. Classical control acts on a point state and a fixed model. On a conservation graph the state is a posterior, a best estimate with its covariance, and the controller is a decision over it: choose the actuator's setting so that the quantity to be protected stays inside its limit with a stated probability, while the rows that estimate the state also say how the actuator moves it.

Stated that broadly the claim is too strong, and the book says where it fails. For a linear plant with a quadratic cost and no constraints, the best move does not depend on the covariance at all. The covariance changes the decision through limits, through losses that are not quadratic, and through moves that also teach. Those are the places this book works.

The book states control as one problem on the belief, the posterior of the state, and each later part works under a restriction of that problem.

The plant grows with the book. Part I works on one equipment hall, a state of two and then three entries, so that every object of the method can be seen whole. Part II draws the hall's discharge loop, with its coil, valve and controller, and later parts add a second unit, a second hall, storage, and an electrical connection.

Every chapter that answers a question follows the same progression, and the progression is the method.

**Step 0: the plant graph.** The plant and its loops, before any question is asked: its nodes, paths and sources; which quantities are pinned and which are free; the physical relations and the controller's rows. The free quantities form the state vector $\mathbf{z}$.

**Step 1: the question.** The control engineer's question in their own words, then as a probability or an expectation of a function of the state.

**Step 2: the graph extended for the question.** A set point or an actuator pinned at a value not yet applied; copies of the graph across time; candidate faults; and the records attached as rows.

**Step 3: the posterior of the state.** Derived and computed once, before any question is answered: the MAP state, the information, the covariance, and the exact posterior where one exists.

**Step 4: the answers.** Each question read from that posterior down to its number.

**Step 5: evaluation and checks.** The gaps between the reported values and the exact values or draws, and, because the records are constructed from a stated truth, the answers against that truth.

**Step 6: the report.** The setting, the probability or the action the engineer acts on, with the exact value beside it where one exists.

**Step 7: the next measurement.** If the band is too wide, the measurements or test signals that could be added, ranked by the width each would remove from the answer that matters.

The framework the book teaches is physics-constrained probabilistic modeling: a plant written as a graph, its unknowns as a state vector, every reading as a row, and every question as a probability over that state. TERRA is the author's own implementation of this framework; the book names the framework and the computations, not the software, and every number in it is reproducible from the derivations as written. The method and its notation are those of the companion books *Reliability Engineering by Derivation* and *Operations and Control by Derivation*, which ask the reliability engineer's and the operator's questions of the same hall.

Each chapter closes with a statement of what to retain, a table of its terms, and exercises at three levels, formulation, derivation and application, each followed by its answer.

# Contents

**How to read this book** **ii**

**Notation and vocabulary** **vii**

**I Thinking in Probabilities** **1**

**1 Controlled plants as graphs** **2**
1.1 One system and several questions . . . 2
1.2 The posterior of the state (step 3) . . . 15
1.3 The answers (step 4) . . . 21
1.4 The checks (step 5) . . . 27
1.5 The report (step 6) . . . 30
1.6 The next measurement (step 7) . . . 32
1.7 What to retain . . . 36
1.8 Exercises . . . 38

**2 The language of control questions** **40**
2.1 Control questions and their mathematical form . . . 40
2.2 The grammar of probability statements . . . 43
2.3 The general forms and the five computations . . . 46
2.4 Estimation, prediction and joint events . . . 57
2.5 Model comparison, diagnosis and decisions . . . 69
2.6 Measurements, classification and workflow . . . 72
2.7 What to retain . . . 79
2.8 Exercises . . . 81

**3 Uncertainty and conditional probability** **84**
3.1 Uncertainty and evidence in a control question . . . 84
3.2 Conditional probability and Bayes' theorem . . . 89
3.3 Bayesian diagnosis of a cooling fault . . . 92
3.4 Marginalization and dependence . . . 94
3.5 Conditioning in a physical model . . . 99
3.6 Predictive variance and the posterior covariance . . . 107
3.7 What to retain . . . 114
3.8 Exercises . . . 115

**4 The anatomy of probabilistic inference** **118**
4.1 From a plant graph to a release probability . . . 118
4.2 The posterior of the state (step 3) . . . 126
4.3 The answers (step 4) . . . 139

4.4 The checks (step 5) . . . 148
4.5 The report (step 6) . . . 153
4.6 The next reading (step 7) . . . 154
4.7 What to retain . . . 158
4.8 Exercises . . . 162

**II The Closed-Loop Problem 164**

**5 Actuators and controllers as graph rows 165**
5.1 The loop drawn in the graph . . . 165
5.2 The posterior of the state (step 3) . . . 170
5.3 The answers (step 4) . . . 171
5.4 The checks (step 5) . . . 177
5.5 The report (step 6) . . . 180
5.6 The next measurement (step 7) . . . 181
5.7 What to retain . . . 183
5.8 Exercises . . . 184

**6 Protected quantities and chance constraints 185**
6.1 A limit the loop's error cannot see . . . 185
6.2 The posterior of the state (step 3) . . . 189
6.3 The answers (step 4) . . . 191
6.4 The checks (step 5) . . . 198
6.5 The report (step 6) . . . 202
6.6 The next measurement (step 7) . . . 203
6.7 What to retain . . . 206
6.8 Exercises . . . 207

**7 The control problem on the graph 209**
7.1 One problem under every controller . . . 209
7.2 The posterior of the state (step 3) . . . 214
7.3 The answers (step 4) . . . 215
7.4 The checks (step 5) . . . 225
7.5 The report (step 6) . . . 226
7.6 The next measurement (step 7) . . . 227
7.7 What to retain . . . 228
7.8 Exercises . . . 229

**8 Stability under parameter uncertainty 231**
8.1 A faster loop on one number . . . 231
8.2 The posterior of the state (step 3) . . . 236
8.3 The answers (step 4) . . . 237
8.4 The checks (step 5) . . . 247
8.5 The report (step 6) . . . 249
8.6 The next measurement (step 7) . . . 249
8.7 What to retain . . . 254
8.8 Exercises . . . 254

**III The Belief** **256**

**9 The Kalman filter as a time-stacked fit** **257**
9.1 A filter block on the hall sensor . . . 257
9.2 The posterior of the state (step 3) . . . 261
9.3 The answers (step 4) . . . 264
9.4 The checks (step 5) . . . 274
9.5 The report (step 6) . . . 275
9.6 The next measurement (step 7) . . . 276
9.7 What to retain . . . 278
9.8 Exercises . . . 279

**10 Feedback as conditioning** **280**
10.1 One new reading and the state the loop acts on . . . 280
10.2 The posterior of the state (step 3) . . . 283
10.3 The answers (step 4) . . . 284
10.4 The checks (step 5) . . . 296
10.5 The report (step 6) . . . 297
10.6 The next measurement (step 7) . . . 299
10.7 What to retain . . . 301
10.8 Exercises . . . 302

**11 Observability and sensor loss** **303**
11.1 A loop that holds a reading it may lose . . . 303
11.2 The posterior of the state (step 3) . . . 308
11.3 The answers (step 4) . . . 310
11.4 The checks (step 5) . . . 321
11.5 The report (step 6) . . . 322
11.6 The next measurement (step 7) . . . 324
11.7 What to retain . . . 326
11.8 Exercises . . . 327

**A Atlas of control questions** **329**

# Notation and vocabulary

Every symbol below has the same meaning in every chapter, except where a chapter's table says otherwise. The unknown always stands on the left of a conditioning bar and the evidence on the right. Vectors and matrices are bold. Symbols of the larger plants are added here as the chapters that introduce them land.

| Symbol | Unit | Meaning |
|---|---|---|
| *The plant graph and its state* | | |
| $\mathbf{z}$, $z_i$ | varies | State vector of the plant graph: every free quantity of one telemetry window at once; its $i$-th entry. |
| $\mathbf{z}_k$, $k$ | varies, – | The state at controller step $k$; the step index, each step 30 s on the closed-loop hall. |
| $\mathbf{y}$, $y$ | varies | The evidence: all readings and prior information treated as known; a single reading, subscripted by what it reads ($y_{T_s}$, $y_{Q_A}$). |
| $r_k(\mathbf{z})$ | varies | Row $k$ of the model: a physical relation that must hold ($r_k = 0$), a reading compared with the state, or a prior. |
| $w_k$, $\sigma_k$ | varies | Weight (precision) of row $k$ and the standard deviation it corresponds to, $w_k = 1/\sigma_k^2$. A row may equivalently be written normalized, $r_k/\sigma_k$, with unit weight. |
| $\sigma_p$ | varies | Tolerance of a physics row; an exact constraint is the limit $\sigma_p \to 0$. |
| *The equipment hall of Part I* | | |
| $T$, $T_c$ | °C | Temperature of the hall air; the supply air held by the cooling unit's discharge controller, pinned. |
| $Q$, $G$ | kW, kW/K | Heat load of the hall; effective conductance of its air path to the supply air. |
| $g$, $G_{\text{ref}}$ | –, kW/K | Log-conductance $\ln(G/G_{\text{ref}})$, the entry of the state that carries the conductance; its reference value. |
| $y_T$, $\sigma_T$ | °C, K | The hall sensor's reading and the standard deviation of its error. |
| *The closed-loop hall of Part II* | | |
| $r$, $e$ | °C, K | The set point of the discharge loop; the control error $e = y_c - r$ on the discharge sensor's reading. |
| $c$, $u$ | – | The valve command of the PI controller; the valve position, from 0 to 1, between its stops. |
| $K_p$, $K_i$ | 1/K, 1/(K s) | Proportional and integral gains. |
| $K_Q$ | K/kW | Process gain of the hall to its load, $\partial T/\partial Q = 1/G$. |

| Symbol | Unit | Meaning |
|---|---|---|
| $W$, $g_w$, $T_w$ | kW/K, –, °C | Coil conductance to chilled water with the valve fully open; its log-derate; the chilled-water temperature, pinned. |
| $C_h$, $C_c$, $\tau_a$ | kJ/K, kJ/K, s | Thermal capacity of the hall air; of the coil and supply air; the actuator's time constant. |
| $T_{\max}$ | °C | A limit on a temperature. |
| $f^*$, $G^*$, $T^*$ | varies | A requirement level on a state entry or a derived quantity; a starred plant quantity is also a scenario value, the quantity at a set point not yet applied. |
| $b_j$ | varies | Bias of reading $j$, an entry freed when a sensor is suspected. |
| *Probability* | | |
| $\Pr(A)$, $\Pr(A \mid B)$ | – | Probability of event $A$; conditional on event $B$. |
| $p(\cdot)$ | varies | A probability density; $p(\mathbf{z} \mid \mathbf{y})$ is the posterior of the state, $p(\mathbf{y} \mid \mathbf{z})$ the likelihood, $p(\mathbf{z})$ the prior. |
| $\mathbb{E}$, Var, Cov | varies | Expectation, variance, covariance. |
| $\mathcal{N}(\mu, \sigma^2)$ | – | Normal distribution with mean $\mu$ and variance $\sigma^2$; $\Phi$ its standard cumulative distribution function, $\varphi$ its density; $\Phi_2(u_1, u_2; \varrho)$ that of two standard normals of correlation $\varrho$. |
| $\chi^2_\nu$ | – | Chi-square distribution with $\nu$ degrees of freedom. |
| *Inference objects* | | |
| $\mathcal{L}(\mathbf{z})$ | – | The objective, minus the log posterior up to a constant: $\mathcal{L} = \frac{1}{2} r(\mathbf{z})^\mathsf{T} r(\mathbf{z})$ for normalized Gaussian rows. |
| $\hat{\mathbf{z}}$ | varies | The maximum a posteriori (MAP) state, the minimizer of $\mathcal{L}$. |
| $\mathbf{J}$ | varies | Jacobian of the normalized rows with respect to the state, evaluated at $\hat{\mathbf{z}}$. |
| $\mathbf{\Lambda}$ | varies | Information matrix (Gauss–Newton Hessian of $\mathcal{L}$), $\mathbf{J}^\mathsf{T}\mathbf{J}$ for normalized rows, $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$ with $\mathbf{W} = \operatorname{diag}(w_k)$ for unnormalized ones. |
| $\mathbf{\Sigma}$, $\Sigma_{ij}$ | varies | Posterior covariance of the state under the Laplace approximation, $\mathbf{\Lambda}^{-1}$; its entries. The posterior spread of entry $i$ is written $\sqrt{\Sigma_{ii}}$; $\sigma_k$ always names a row's standard deviation. |
| $f(\mathbf{z})$, $\mathbf{g}$ | varies | A derived quantity of the state and its gradient $\partial f / \partial \mathbf{z}$ at $\hat{\mathbf{z}}$. |
| $\mathbf{A}$, $\mathbf{A}_{\text{free}}$, $\mathbf{A}_{\text{dep}}$ | varies | Jacobian of the physics rows on the full state; its blocks on the free and the dependent entries (Chapter 4). |
| $\mathbf{N}$, $\mathbf{J}_{\text{rows}}$, $\mathbf{J}_{\text{red}}$ | varies | Null-space basis of the physics rows, $[\mathbf{I};\ -\mathbf{A}_{\text{dep}}^{-1}\mathbf{A}_{\text{free}}]$; the Jacobian of the evidence rows (readings and priors); the reduced Jacobian $\mathbf{J}_{\text{rows}}\mathbf{N}$ of the reduced state. |
| $f_L$, $f_U$ | varies | The ends of the credible interval of a derived quantity. |
| $y_{\text{new}}$, $\tilde{y}_a$ | varies | A future reading (the predictive) and the outcome of a candidate reading $a$ before it is taken (the preposterior). |
| $e_k$ | – | Normalized residual of reading $k$: its residual divided by the standard deviation of that residual. |

| Symbol | Unit | Meaning |
|---|---|---|
| $M$, $M_j$, $B_{AB}$ | – | A drawing (model) of the plant; the members of a family of candidates; the Bayes factor $p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$. |
| $p(\mathbf{y} \mid M)$ | – | Evidence (marginal likelihood) of model $M$. |
| $\mathbf{J}_a$, $\mathbf{\Lambda}_a$, $\mathbf{\Sigma}_a$ | varies | The Jacobian rows of a candidate reading, its information $\mathbf{J}_a^\mathsf{T}\mathbf{J}_a$, and the preposterior covariance $(\mathbf{\Sigma}^{-1} + \mathbf{\Lambda}_a)^{-1}$. |
| $\gamma$, $t^*$ | varies/h, h | Slope of an entry's trend across windows; the time at which a rising trend reaches a limit from below (as in the companion book). |
| *Decisions* | | |
| $\alpha$, $\kappa$ | – | The stated risk of a chance constraint, $\Pr(f \le f^* \mid \mathbf{y}) \ge 1 - \alpha$; the stated confidence $\kappa = 1 - \alpha$. |
| $U(a)$, $L(a, \cdot)$ | varies | Preposterior value of a candidate reading $a$; the loss of an action $a$. |

## Vocabulary

*Plant graph* The physical system as nodes (places where a conserved quantity balances), edges (paths along which it flows) and sources, with the relations that must hold among them. Built before any question is asked.

*State vector* The collection of free quantities the fit estimates at once; outputs that are entries of the state need no further transformation.

*Pinned, free* A quantity treated as known (boundary condition) against one carried in the state.

*Residual, row* A relation written so that it equals zero when satisfied; physics rows constrain the state, reading rows compare it with telemetry, prior rows carry handbook or commissioning knowledge.

*State estimation* The fit of the state to its readings and priors: the MAP state and its covariance.

*Observability* Whether the readings determine an entry of the state, or only a combination of entries; read from the Jacobian and from the posterior spread against the prior spread.

*Set point, process gain* The value a loop is asked to hold; the change of the controlled quantity per unit change of an input.

*Chance constraint* A limit required to hold with a stated probability, $\Pr(f \le f^* \mid \mathbf{y}) \ge 1 - \alpha$.

*Back-off* The distance a set point stands inside its limit because the quantity is uncertain: $\Phi^{-1}(1 - \alpha)$ times the quantity's spread, for a normal posterior.

*Scenario node* The plant re-solved with a set point or an actuator pinned to a value not yet applied, and the covariance carried through.

*Laplace approximation* Replacing the posterior by the Gaussian whose covariance is the inverse curvature of the objective at the MAP state.

*Delta method* Linearizing a derived quantity around the MAP state so that its variance is $\mathbf{g}^\mathsf{T}\mathbf{\Sigma}\mathbf{g}$.

*Posterior predictive* The distribution of a future reading obtained by averaging its model over the posterior.

*Evidence, Bayes factor* The normalizing constant of one model's posterior; the ratio of two models' evidences.

*Preposterior* The posterior that would result from a reading not yet taken; the object that values a reading before it is requested.

*The five computations* Posterior, trend, evidence, preposterior and expected consequence: the objects every control question is answered from.

# Part I

# Thinking in Probabilities

# Chapter 1

# Controlled Plants as Graphs

## 1.1 One system and several questions

A commissioning engineer has the trend of a data center hall's discharge loop on one screen, flat at 18.0°C, and on another the one hall sensor, with an error of about half a kelvin, at 23.2°C with 30 kW on the load meter. The owner asks to double the load to 60 kW, with the hall air below 27.5°C. The commissioning sheet takes the reading as exact. The hall stands 5.2 K above the supply air under 30 kW, a process gain of 0.1733 K/kW, so 60 kW runs the hall at 28.4°C, 0.9 K over the limit. The sheet's conclusion is one line: lower the set point 0.9 K to 17.1°C and the hall holds.

The design handbook says otherwise. It gives the conductance of the hall's air path to the supply air as 7.5 kW/K, right to within a factor of about 1.35, which puts the hall at 26°C under 60 kW, with nothing to do. Each number treats its own input as exact, and neither is weighed against the other. The sheet cannot say with what probability 17.1°C holds the hall, or whether the coil can deliver it. A flat loop trend says the loop holds its reading. It says nothing about the hall. Answering needs the handbook value, the reading and the energy balance that ties them together in one description of the hall, each with its own uncertainty, and the set point as a value of that description the engineer can move; the sibling volume *Operations and Control by Derivation* reads this same hall with the supply air as a fixed boundary, and here the set point can be moved; Part II draws the loop. The question of this chapter is therefore

*how a controlled plant is represented so that one representation serves every question asked of it (its state, its gains, its set point, the next test), and what must be fixed before any inference is performed.*

**Jargon: plant graph, node, path, source, pinned and free quantities, set point, discharge loop, process gain, state vector, residual, row, observation, prior, posterior, state estimation, model discrepancy, scenario node, Jacobian, normal equations, information matrix, covariance, Laplace approximation, factor graph, commissioning log**

The *discharge loop* is the controller that holds the supply air at the value it is asked to hold, its *set point*. A *process gain* is the steady change of an output per unit change of an input; here the input is the load, a disturbance, so this is strictly the disturbance gain. The *commissioning log* is the engineer's dated record of what was found, what was decided and what is to be measured next. A *plant graph* represents a physical system as *nodes*, places where a conserved quantity such as energy or mass must balance; *paths* (edges), along which the quantity flows; and *sources*, which supply it. A quantity is *pinned* when it is treated as known, as a boundary condition, and *free* when it is unknown and must be inferred. The free quantities, collected in one vector, form the *state vector* $\mathbf{z}$. A *residual* is a relation written so that it equals zero when it is satisfied; each residual is a *row* of the model. A physical row states a relation that must hold among entries of the state; an *observation* (a reading) is a measured value related to an entry of the state with an error, and its row compares the two. The *prior* expresses what is known about the state before the readings, for example a handbook value of a conductance; the *posterior* $p(\mathbf{z} \mid \mathbf{y})$ expresses what is known after the evidence $\mathbf{y}$; fitting the state to its rows, which yields the posterior's most probable state and its spread, is what the field calls *state estimation*. A *model discrepancy* is the difference between the physical system and the model of it that would remain if every entry of the state were known exactly. A *scenario node* is a node added to the graph for an operating condition that has not occurred, so that a question can be asked about it. The *Jacobian* $\mathbf{J}$ is the matrix of derivatives of every row with respect to every entry of the state, one row per row of the model and one column per entry. The *normal equations* $(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$ give the step that best reduces all the rows at once when each is replaced by its tangent. The *information matrix* $\mathbf{J}^\mathsf{T}\mathbf{J}$ states how sharply the rows together fix the state, direction by direction, and its inverse is the posterior *covariance* $\boldsymbol{\Sigma}$, whose diagonal holds the variances of the entries and whose off-diagonal entries hold how they vary together. The *Laplace approximation* replaces the posterior by the normal distribution centered on its most probable state with that covariance. A *factor graph* draws the entries of the state and the rows that act on them, one node each.

**Key idea**

A graph describes the controlled plant; a probabilistic question states what is to be learned about it. The plant graph, its pinned and free quantities and its physical rows are built once, and each question is a probability over the state of that one graph, with the unknown on the left of the conditioning bar and the evidence on the right. The posterior of the state is carried by the most probable state $\hat{\mathbf{z}}$, the covariance $\boldsymbol{\Sigma}$ and the Jacobian $\mathbf{J}$ of the rows, and every answer is an entry of the state read off them or a function of the state carried through them.

### 1.1.1 Several questions about one equipment hall

A commissioning engineer responsible for the hall might ask any of the following:

1. What temperature is the hall running at now?
2. What is the process gain to the load?
3. Can the hall carry 60 kW within its limit at the present set point?
4. Can the set point fix it, and how far must it move?
5. Would another sensor sharpen the gain?

6. What happens if the path loses half its conductance at once?

The questions are different. The first two ask for the value of an unknown quantity; the third asks for the probability that a future temperature stays within a limit; the fourth asks the same probability at a set point not yet applied; the fifth asks what a reading not yet taken would be worth; the sixth asks about a transient that has not occurred. All six, however, concern the same physical system: the same hall, the same air path, the same sensor.

Constructing a separate mathematical model for every question would repeat the same physics six times, and the six models could disagree about it. The method of this book begins instead with one common representation of the equipment and its physical relations, the *plant graph*. It is step 0 of the method stated in *How to read this book*: the system as it is, before any question is asked, built once and kept for every question that follows. Section 1.1.6 writes the questions that this chapter answers over the graph.

### 1.1.2 The system, the question and the evidence

Three things must be fixed before any probabilistic inference is performed, and the plant graph is where they are fixed. The first is the division of the quantities: which are known and pinned (the supply-air temperature, the heat load), which are unknown and free (the hall temperature, the conductance), and which are observed (the sensor reading). The field often separates the unknowns into *states*, such as temperatures and flows, and *parameters*, such as conductances and rates; in this book both kinds are entries of one state vector, and the inference treats them alike. The second is the set of physical relations that must hold among the quantities, such as conservation of energy at the hall. The third is the evidence: the readings, the records and the prior knowledge, attached to the graph as rows of their own.

The question is a separate object. It is stated over the graph after the graph is built, and it decides what is computed, not what the system is. The separation makes the questions of a commissioning systematic: when a new question arrives, the graph is already in hand, and the work is to write the question over its state.

> **Principle 1.1.** *The physical model is reusable; the probabilistic question determines what is calculated from it. A plant graph is built once from the equipment and its physics, and every question is a probability or an expectation over the state of that graph.*

### 1.1.3 Larger plant graphs and factor graphs

The equipment hall of this chapter has one thermal path and two unknown quantities, the smallest graph with every kind of element. The same construction serves a plant of many components, and the graph on which an inference operates is a second object, drawn from the plant graph and the evidence attached to it.

**A cooling network.** Real plants may include hundreds or thousands of components, each with its own equations and uncertainties. Consider a cooling network with two chillers, two pumps, a common electrical supply and several temperature sensors. Its plant graph may contain temperatures, pressures, flows and electrical states at its nodes; physical components represented by paths or by component relations; heat, mass and power conservation equations; telemetry and historical records; the run and trip states of equipment; and shared parameters and shared dependencies.

Figure 1.1 draws a reduced version. The green boxes are free nodes: the two chillers and the two pumps, each carrying its status and capacity or flow, the supply header with its water temperature $T_s$ and flow, and the hall with its temperature $T$. The gray box on the left is the pinned outdoor air, to which the chillers reject heat; the white box in the middle is the common

electrical supply, which feeds all four machines; the white box at the upper right is the heat load of the hall. Solid green arrows are physical paths of heat and water; thin gray arrows are electrical feeds; dashed arrows are observations, two temperature readings and the maintenance record of a chiller. The common supply is a single node on which four components depend: if it trips, all four stop together, and a graph that omitted it would treat their trips as independent.

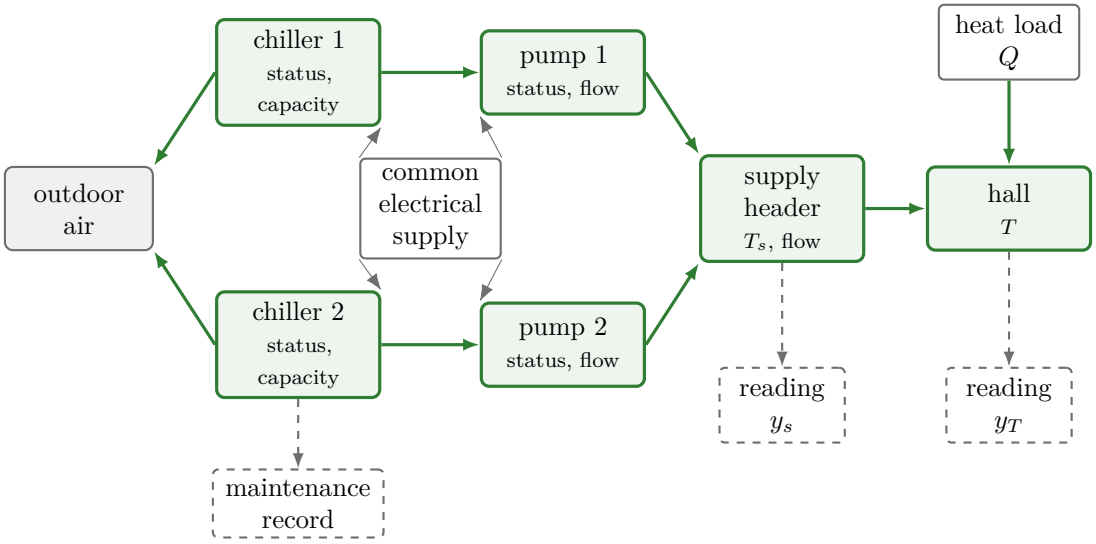



Figure 1.1: Plant graph of a reduced cooling network (illustrative; return piping omitted). Free: the status and capacity of each chiller, the status and flow of each pump, the header water temperature $T_s$ and flow, and the hall temperature $T$. Pinned: the outdoor air, to which the chillers reject heat, and the heat load $Q$; the common electrical supply feeds all four machines and is the shared dependency. Observed: $T_s$ and $T$ through two readings, and chiller 2 through its maintenance record.

The mathematical principle is the one the single hall shows below. The network is a collection of physical relations,

$$r_1(\mathbf{z}) = 0, \quad r_2(\mathbf{z}) = 0, \quad \ldots, \quad r_m(\mathbf{z}) = 0,$$

one per balance and per component relation, over a state vector $\mathbf{z}$ that now has many entries. Observations and probability models for the uncertain quantities are attached as rows of their own. What is ultimately sought is the joint posterior of the state,

$$p(\mathbf{z} \mid \mathbf{y}, M),$$

where $M$ represents the assumed system model, including its structure: which components exist, how they are connected and which dependencies are represented. Writing $M$ explicitly matters because the structure itself may be uncertain, and a question about the structure is answered by comparing two such models, not by the posterior inside either.

**Physical graphs and factor graphs.** A physical graph and a probabilistic factor graph are related, but they are not the same object. The physical graph represents equipment and connections: halls, paths, chillers, pumps, sensors. A *factor graph* represents the variables of the inference and the mathematical relations among them, including the observations, the priors and the dependencies: each entry of the state is one node, each row of the model is one factor node, and each factor is linked to the entries its row involves. It is constructed from the physical graph and the evidence attached to it, and it shows which row acts on which unknown. The distinction becomes important when competing system structures and diagnosis are studied in later chapters. Two physical drawings of the same plant, one with a single air path and

one with two, give two factor graphs with different variables and rows, and comparing them is comparing two models $M$. A sensor bias adds a variable to the factor graph without changing the equipment at all.

### 1.1.4 The plant graph (step 0)

The simplest plant graph with every kind of element has one free node, one pinned node, one path, one source and one observation, and the equipment hall is that graph. It is built before any question is asked and before any number is fitted.

**Nodes, a path, a source and an observation.** Consider a simplified equipment hall with one heat-removal path, the hall's air path to the supply air that the cooling unit delivers. Figure 1.2 draws it as a plant graph. The green box on the left is the hall, a free node whose air temperature $T$ is unknown; the gray box on the right is the supply air, a pinned node whose temperature $T_c = 18°\text{C}$ is held by the discharge loop. The solid green arrow between them is the heat-removal path, characterized by its effective thermal conductance $G$. The box above the hall is the heat load $Q$ supplied by the servers, a source. The dashed box below is the hall sensor, and the dashed arrow from the hall to it is an observation: the sensor reports a value related to $T$, not $T$ itself. Solid connections are physical relations; the dashed connection is evidence.

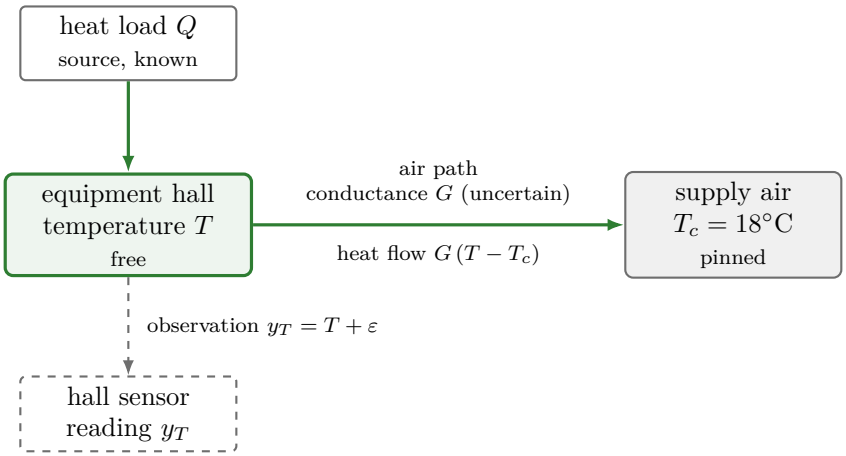



Figure 1.2: Plant graph of the equipment hall in steady state. Free: the hall air temperature $T$ and the conductance $G$ of the air path. Pinned: the supply-air temperature $T_c = 18°\text{C}$, the set point of the discharge loop, and the heat load $Q$. Observed: $T$, through the reading $y_T$ with error $\varepsilon$. Solid arrows are physical relations; the dashed arrow is an observation.

Table 1.1 lists the elements of the graph. The first column names each element, the second its role in the graph, the third its symbol and the fourth its unit; the last column states how the element enters this first example.

For now the supply-air temperature and the heat load are assumed known, and the effective conductance is uncertain. The supply air $T_c$ is the discharge loop's set point $r = 18°\text{C}$, held by a controller that Part I treats as ideal: this is what makes the supply air a pinned boundary, with no spread, and it fails when the valve reaches its stop (Section 1.6.5 and Chapter 15) and if the discharge sensor is off (Chapter 6). The conductance lumps every mechanism by which the hall's heat reaches the supply air, the airflow through the cooling unit and the mixing of the hall air alike, into one number; how well that lumping represents the hall is a modeling assumption, examined in Section 1.6.5. This is the starting physical model.

Table 1.1: The elements of the plant graph of Figure 1.2.

| Element | Role | Symbol | Unit | In the example |
|---|---|---|---|---|
| Equipment hall | free node | $T$ | °C | air temperature unknown |
| Supply air | pinned node | $T_c$ | °C | held at the set point $r = 18$°C by the discharge loop |
| Heat load | source | $Q$ | kW | thermal power supplied by the servers, known |
| Air path | path (edge) | $G$ | kW/K | effective conductance, uncertain |
| Hall sensor | observation | $y_T$ | °C | a reading related to the hall temperature |

**Conservation of energy at the hall node.** A graph becomes useful for computation when the mathematical relations between its elements are specified. Assume steady state: the hall stores no more and no less heat from one moment to the next, so that its temperature is constant. Conservation of energy at the hall node then requires the heat entering to equal the heat leaving,

$$Q_{\text{in}} = Q_{\text{out}},$$

where $Q_{\text{in}}$ and $Q_{\text{out}}$ are thermal powers in kilowatts. The heat entering is the load, $Q_{\text{in}} = Q$. The heat leaving is modeled as proportional to the temperature difference across the path,

$$Q_{\text{out}} = G\,(T - T_c),$$

the defining relation of the effective conductance: $G$ is the heat flow per kelvin of difference between the hall and the supply air. Combining the two,

$$Q = G\,(T - T_c), \tag{1.1}$$

where $Q$ is the heat generated by the servers in kW, $G$ the effective thermal conductance in kW/K, $T$ the hall air temperature in °C and $T_c$ the supply-air temperature in °C. A temperature difference has the same value in kelvin and in degrees Celsius, so the units balance: kW/K times K is kW.

Equation (1.1) is a *physical relation*, not a probability statement. It states what must hold if the simplified thermal model is correct, whatever is or is not known about its quantities. Since $G > 0$, dividing by $G$ and adding $T_c$ gives

$$T = T_c + \frac{Q}{G}, \tag{1.2}$$

which gives the temperature when the other three quantities are specified. The hall is warmer than the supply air by $Q/G$: more load or less conductance raises it.

**The forward calculation.** The hall is operated at the load $Q = 30\,\text{kW}$, and its air path was designed for a conductance of $G = 6\,\text{kW/K}$. With $T_c = 18$°C, equation (1.2) gives

$$T = 18 + \frac{30\,\text{kW}}{6\,\text{kW/K}} = 18 + 5 = 23\text{°C},$$

the steady-state temperature of the hall at its design conductance. The calculation used no probability, estimated no uncertainty and used no observational evidence. It is a deterministic *forward calculation*: from pinned inputs and a specified conductance to the temperature they imply. Every later chapter needs it, because every inference compares temperatures computed this way with what the sensor reports. Doubling the load doubles the excess $T - T_c$ at every conductance: at $G = 6\,\text{kW/K}$ the hall at 60 kW would be at 28°C.

**The state vector.** The forward calculation required the conductance to be known. In the inference problems of this book it is not, and neither is the hall temperature. The *state vector* is the collection of the unknown quantities that the inference estimates at once. In the equipment hall the unknowns are the hall temperature and the thermal conductance, and one possible representation is

$$\mathbf{z} = \begin{bmatrix} T \\ G \end{bmatrix}.$$

The state holds everything the inference treats as free; the pinned quantities $T_c$ and $Q$ are not in it, and neither is the reading $y_T$, which is evidence. The concept recurs throughout the book: the graph fixes the state, the evidence constrains it, and each question is a statement about it. Quantities that are entries of the state, such as $T$ here, are answered directly by the inference; quantities that are functions of the state, such as the heat flow through the path, require one further transformation.

**The log-conductance.** A conductance is positive, and an estimate of it must be positive too. In practice the conductance is therefore often represented by a logarithmic entry,

$$g = \ln\!\left(\frac{G}{G_{\mathrm{ref}}}\right),$$

where $G_{\mathrm{ref}} > 0$ is a reference conductance in kW/K and $g$ is dimensionless. Inverting,

$$G = G_{\mathrm{ref}}\, \mathrm{e}^{g},$$

which is positive for every real $g$: an inference free to move $g$ anywhere on the real line can never produce a negative or zero conductance. A change in $g$ is a relative change in $G$, since $\mathrm{d}G = G_{\mathrm{ref}}\, \mathrm{e}^{g}\, \mathrm{d}g = G\, \mathrm{d}g$, so that $\mathrm{d}g = \mathrm{d}G/G$; uncertainty in $g$ reads directly as a fractional uncertainty in $G$. The reference value only sets the origin of $g$. In this chapter it is taken as $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$, the design conductance of the forward calculation, so that $g = 0$ there, $G = 7.5\,\mathrm{kW/K}$ corresponds to $g = \ln 1.25 = 0.2231$. The state vector becomes

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} T \\ g \end{bmatrix}, \tag{1.3}$$

with entries $z_1 = T$ in °C and $z_2 = g$, dimensionless. It is the one state of this chapter: every question below is a statement about these two entries.

**The physics row.** The thermal relation can now be written in terms of the state. Substituting $G = G_{\mathrm{ref}}\, \mathrm{e}^{g}$ into equation (1.1),

$$Q = G_{\mathrm{ref}}\, \mathrm{e}^{g}\, (T - T_c),$$

or, moving everything to one side,

$$r_{\mathrm{phys}}(\mathbf{z}) = G_{\mathrm{ref}}\, \mathrm{e}^{g}\, (T - T_c) - Q = 0. \tag{1.4}$$

The function $r_{\mathrm{phys}}(\mathbf{z})$ is a *residual*, in kW. A residual measures how far a candidate state is from satisfying a relation: it is zero when the relation holds exactly and grows as the candidate departs from it. With $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$, $T_c = 18°\mathrm{C}$ and $Q = 30\,\mathrm{kW}$, the candidate $\mathbf{z} = (23°\mathrm{C},\ 0)$ gives $r_{\mathrm{phys}} = 6 \times 1 \times 5 - 30 = 0$: it is the forward calculation. The candidate $\mathbf{z} = (23.2°\mathrm{C},\ 0)$ gives $6 \times 5.2 - 30 = 1.2\,\mathrm{kW}$: at that temperature and conductance the path would remove 1.2 kW more than the servers supply, so the hall could not be in steady state. The candidate $(23.2°\mathrm{C},\ \ln(30/31.2)) = (23.2°\mathrm{C},\ -0.0392)$ again gives zero: the temperature 23.2°C is consistent with the physics if the conductance is $30/5.2 = 5.769\,\mathrm{kW/K}$. The zero set

of $r_{\text{phys}}$ is the curve $T = T_c + Q/(G_{\text{ref}}\,\mathrm{e}^{g})$ in the plane of $(g, T)$; the physics alone allows every point of that curve and no point off it.

The same row has a second form, in the units of the quantity it fixes. Dividing equation (1.4) by $G_{\text{ref}}\,\mathrm{e}^{g} = G$, which is positive, leaves the zero set unchanged and gives the row in kelvin, the departure of a candidate temperature from the temperature the physics implies. Divided further by a *model tolerance* $\sigma_p$, the temperature error in kelvin that the physics is allowed, it becomes a dimensionless row,

$$r_1(\mathbf{z}) = \frac{T - T_c - (Q/G_{\text{ref}})\,\mathrm{e}^{-g}}{\sigma_p}, \tag{1.5}$$

which counts the violation of the physics in multiples of its tolerance. This is the form in which the physics row enters every computation of the book. The formulation extends to systems of any size. A larger physical system is represented by a collection of residuals, one per relation: an electrical network has rows for Kirchhoff's current and voltage laws, a fluid network rows for mass conservation at its junctions and for the pressure losses along its pipes, and a cooling plant rows for the energy balance at each of its nodes. Each row has the form of equation (1.4), and the inference treats them alike.

**Exact constraints and model discrepancy.** Physical rows and reading rows should not automatically be treated as equivalent kinds of noisy measurement. Conservation of energy is a law: if the hall is in steady state and the model has captured every path, equation (1.4) holds exactly, and it can be imposed as an *exact constraint*, $r_{\text{phys}}(\mathbf{z}) = 0$, which restricts the state to the zero set of the row. In a computation the exact constraint is carried as the physics row (1.5) with a model tolerance $\sigma_p$ far smaller than every other spread of the problem, and exactness is the limit $\sigma_p \to 0$; the limit is checked by showing that the results stop moving as $\sigma_p$ shrinks, never by using the row to eliminate an entry of the state. A reading row, by contrast, is meant to be violated by an amount of the order of the sensor's error: the reading is not the temperature.

Approximate physics is a third case. The single effective conductance of equation (1.1) is an idealization; the real hall has several paths, a conductance that varies with airflow, and stored heat. When the relation is only approximately true, the honest model is

$$G_{\text{ref}}\,\mathrm{e}^{g}\,(T - T_c) - Q = \delta,$$

with an explicit *model discrepancy* $\delta$ in kW, an entry of its own with a stated prior, for example $\delta \sim \mathcal{N}(0, \sigma_\delta^2)$. Algebraically this is a physics row with the weight $1/\sigma_\delta^2$, and in the objective it looks like a reading row. It behaves differently. A measurement error is new at each reading and averages out over many readings; a discrepancy is a property of the model and repeats wherever the same operating condition repeats, so that many readings at one condition do not reduce it. Treating a discrepancy as measurement noise therefore makes the inference overconfident, and treating a law as noisy lets the inference trade physics against data. The choice between an exact constraint and an explicit discrepancy is made when the graph is built, and it is stated with the model. In this chapter conservation of energy is exact.

### 1.1.5 The record and the handbook value

The record of the hall is one reading. The hall sensor, the one temperature sensor installed in the hall, reports at the load of 30 kW

$$y_T = 23.2^{\circ}\mathrm{C}.$$

The sensor is not perfectly accurate. Its measurement model is

$$y_T = T + \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0, \sigma_T^2),$$

where $\varepsilon$ is the measurement error in kelvin, assumed normal with zero mean, and $\sigma_T$ its standard deviation, which the sensor's specification gives as

$$\sigma_T = 0.5\,\mathrm{K}.$$

The reading is the only observation: the supply-air temperature and the load are pinned, and nothing measures the conductance.

Knowledge of the conductance that does not come from the reading is also in hand. A design handbook, the commissioning documents or earlier tests of similar halls give an estimate of the effective conductance with an uncertainty; for this hall (illustrative) the handbook gives 7.5 kW/K, with the true conductance within a factor of about 1.35 of it with probability of about 0.68. The handbook value is not a measurement of this hall. It is information held before the reading, and it enters the model as a prior.

### 1.1.6 The questions (step 1)

The engineer's questions can now be written over the graph, each as a probability or an expectation over the state, with the unknown on the left of the bar and the evidence on the right. Table 1.2 lists the six that this chapter answers: the first column is the question in the engineer's words; the second the general form that every question of its category takes, written in the book's notation for a quantity $f(\mathbf{z})$ derived from the state, a threshold $f^*$ on it and the posterior density $p(\mathbf{z} \mid \mathbf{y})$; the third its mathematical form for the hall; the fourth its category among the five named at the front of the book, which Chapter 2 develops; and the last the subsection where it is answered. The first two are estimation, of an entry of the state and of a quantity derived from one, the process gain $K_Q = \partial T/\partial Q = 1/G$ with $G = G_{\mathrm{ref}}\,\mathrm{e}^{g}$. The next three are predictions under a threshold, tail probabilities of the posterior: the third against the hall's present limit of 24°C, the fourth at a load of 60 kW at which the hall has never run, whose temperature $T_2$ is a quantity the graph acquires in step 2, and the fifth at the same load with the set point moved to a value $r$ not yet applied. The sixth is a design question about a reading not yet taken; it asks nothing new of the present evidence, and it is answered as the next measurement for the first five.

The questions differ in what they compute and agree in what they are computed over: the same two entries of the state, under the same posterior. The last question of the list, the loss of half the path's conductance at once, is not in the table: it is a question about a transient that the steady-state graph cannot represent, and Section 1.6.5 states what the graph would need for it. Also absent is a question the list leaves out, whether the hall sensor is biased. That question compares two complete models of the same reading, the plant with an unbiased sensor and the plant with a bias entry on the reading row, and its answer is a ratio of their evidences rather than a property of either posterior; when every sensor of a larger plant is a candidate, the same comparison run over the candidates is a diagnosis. Chapters 26 and 27 answer it, and the residual check of Section 1.4.3 takes a first, coarse look at it.

### 1.1.7 The evidence as rows (step 2)

Step 2 attaches the reading and the handbook value to the graph as rows on the state $\mathbf{z} = (T, g)^{\mathsf{T}}$, and extends the graph with the operating condition the owner's request names. The three rows, the physics row of Section 1.1.4 among them, are Gaussian, and each is written normalized, with a residual and a Jacobian row.

**The modeling choices.** One choice shapes the physics row. The row carries the balance of Section 1.1.4 in the limit $\sigma_p \to 0$: this is what confines the posterior to the physics curve, so that a reading of the temperature informs the conductance, and it fails if heat leaves the

Table 1.2: The questions asked of the equipment hall, the general form of their category, their mathematical form for the hall, their category and where each is answered. $T_2$ is the hall temperature at the load $Q_2 = 60\,\mathrm{kW}$ and $T_2(r)$ the same at the set point $r$ (Section 1.1.7); $\boldsymbol{\Sigma}_a$ the covariance expected after a candidate row with Jacobian row $\mathbf{J}_a$.

| Control engineer's question | General form | Mathematical form | Category | Answered in |
|---|---|---|---|---|
| What temperature is the hall running at now? | $\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$, $\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$ | $p(T \mid \mathbf{y})$, $T = z_1$ | estimation | 1.3.1 |
| What is the process gain to the load? | $\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$, $\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$ | $p(K_Q \mid \mathbf{y})$, $K_Q = \mathrm{e}^{-z_2}/G_{\mathrm{ref}}$ | estimation of a derived quantity | 1.3.1 |
| Is the hall within its 24°C limit now? | $\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$ | $\Pr(T < 24^\circ\mathrm{C} \mid \mathbf{y})$ | prediction under a threshold | 1.3.2 |
| Can the hall carry 60 kW within 27.5°C at this set point? | $\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$ | $\Pr(T_2 < 27.5^\circ\mathrm{C} \mid \mathbf{y})$, $T_2 = T_c + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-z_2}$ | prediction under a threshold | 1.3.2 |
| How far must the set point move to hold it? | $\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$ at a scenario node | $\Pr(T_2(r) < 27.5^\circ\mathrm{C} \mid \mathbf{y})$, $T_2(r) = r + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-z_2}$ | prediction under a threshold | 1.3.3 |
| Which added sensor would narrow the gain? | preposterior covariance $(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$ | $\boldsymbol{\Sigma}_a$ for each candidate row | design | 1.6.1 |

hall by a mechanism the single conductance does not lump or if the hall is not in steady state (Section 1.6.5).

**The reading row.** The reading of Section 1.1.5 becomes the normalized observation residual

$$r_{\mathrm{obs}}(\mathbf{z}) = \frac{T - y_T}{\sigma_T}, \tag{1.6}$$

which is dimensionless: the difference between the candidate temperature and the reading, in units of the sensor's standard deviation. At $T = 23.2^\circ\mathrm{C}$ it is zero; at the design temperature $T = 23^\circ\mathrm{C}$ it is $(23 - 23.2)/0.5 = -0.4$. The row carries the probability of the reading. Given the state, the reading differs from $T$ only by the error, so its density is the normal density of $\varepsilon$ evaluated at $y_T - T$:

$$p(y_T \mid \mathbf{z}) = \frac{1}{\sqrt{2\pi}\,\sigma_T} \exp\left[-\frac{(y_T - T)^2}{2\sigma_T^2}\right] = \frac{1}{\sqrt{2\pi}\,\sigma_T} \exp\left[-\tfrac{1}{2}\, r_{\mathrm{obs}}(\mathbf{z})^2\right].$$

Read as a function of the state for the fixed reading, this is the *likelihood* of the state. A candidate state with a small $|r_{\mathrm{obs}}|$ makes the reading probable; the ratio of the likelihoods of $T = 23^\circ\mathrm{C}$ and $T = 23.2^\circ\mathrm{C}$ is $\mathrm{e}^{-0.4^2/2} = \mathrm{e}^{-0.08} = 0.923$, so the reading alone hardly distinguishes the two. The likelihood does not depend on $g$: the sensor sees the temperature, and it informs the conductance only through the physics row that ties the two together.

**The prior row.** The handbook value becomes a row of the same form. Written as a normal distribution on the log-conductance, $g \sim \mathcal{N}(g_0, \sigma_g^2)$, it is the row $(g - g_0)/\sigma_g$: a handbook value is treated as a reading of $g$ taken before the hall's own. The handbook conductance of 7.5 kW/K gives $g_0 = \ln(7.5/6) = 0.2231$, and $\sigma_g = 0.3$ states that the true conductance is within a factor

$e^{0.3} = 1.35$ of it with probability of about 0.68. The prior is written $p(g)$ in general; its form is a modeling choice that must be stated and justified, and Section 1.4.4 tests how much this one weighs.

**Three kinds of row.** The model now holds three very different kinds of relation. Table 1.3 sets them side by side. Each row of the table is one kind of relation; the columns give its equation for the hall, its residual and its role.

Table 1.3: The three kinds of row in the equipment hall model, with $G = G_{\text{ref}}\, e^{g}$.

| Row | Equation | Residual | Role |
|---|---|---|---|
| Physics | $G\,(T - T_c) - Q = 0$ | $G_{\text{ref}}\, e^{g}(T - T_c) - Q$ | constrains what is physically possible |
| Reading | $y_T = T + \varepsilon$ | $(T - y_T)/\sigma_T$ | provides imperfect information about the state |
| Prior | $g \sim \mathcal{N}(g_0, \sigma_g^2)$ | $(g - g_0)/\sigma_g$ | carries knowledge held before the readings |

The distinction between the first two rows is foundational. The physical equation constrains the system: it says which combinations of temperature and conductance can occur at all. The reading provides evidence about the unknown quantities: it says which of the possible combinations are plausible in light of what was measured. The prior row has the form of a reading row and the role of knowledge held in advance. A probabilistic inference method combines the three kinds of relation to obtain a distribution over the plausible states.

**Rows and weights.** Chapter 3 shows that, for normal errors, the negative logarithm of the posterior is, up to a constant, the objective $\mathcal{L}(\mathbf{z}) = \frac{1}{2}\sum_k w_k\, r_k(\mathbf{z})^2$, a weighted sum of the squared rows with weights $w_k = 1/\sigma_k^2$. A reading row may be written either normalized, as in equation (1.6), with unit weight, or in kelvin, as $T - y_T$ with weight $1/\sigma_T^2$; both put the same term $(T - y_T)^2/\sigma_T^2$ into the objective. Every row of this chapter is written normalized: the physics row in the form (1.5) with the tolerance $\sigma_p$, the reading row (1.6) and the prior row, so that every row counts in standard deviations and none dominates because of its unit.

**The scenario node.** The owner's request doubles the electrical load from 30 kW to 60 kW, an operating condition at which the hall has never run. A second operating condition is added to the graph, sharing the same unknown conductance. It is a closed-loop scenario node: the set point is pinned and the valve, which Part I does not draw, is left to the loop. Figure 1.3 draws it. The upper row is the present condition, at which the hall has been operated and observed: the free node at $Q_1 = Q = 30\,\text{kW}$ with temperature $T$, its path to the pinned supply air, and its sensor. The lower row is the full-load condition, a scenario node at $Q_2 = 60\,\text{kW}$ with temperature $T_2$ and its own path to the same supply air. The vertical dashed line between the two paths marks that they carry one and the same conductance. The scenario node has no sensor: it is in the graph only because the question asks about it.

The two operating conditions satisfy the energy balance of equation (1.1) each at its own load,

$$Q_1 = G\,(T - T_c), \tag{1.7}$$

$$Q_2 = G\,(T_2 - T_c), \tag{1.8}$$

with $Q_1 = Q$. The two temperatures are different, but the conductance is shared, and that connection is what makes the second condition predictable from the first. The scenario node adds

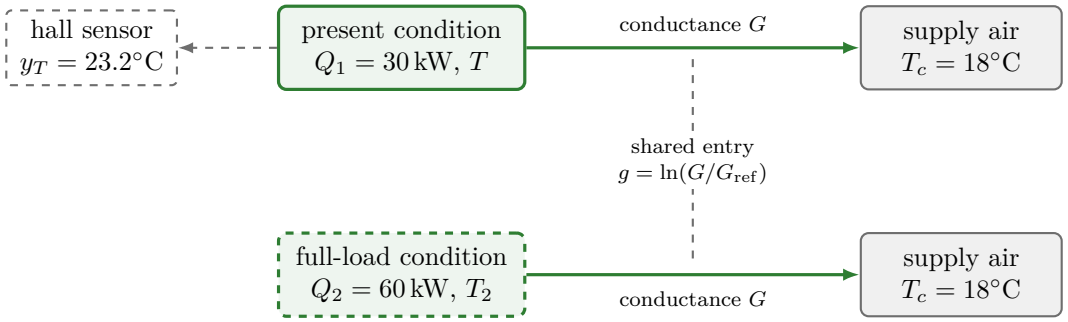



Figure 1.3: Two operating conditions of the equipment hall linked through the same physical conductance. Free: the present-condition temperature $T$ and the shared log-conductance $g$, the two entries of the state; the full-load temperature $T_2$ (dashed node) is a function of $g$. Pinned: the supply-air temperature and the two loads. Observed: $T$ only, through the reading $y_T$; the scenario node is unobserved and enters the graph because the question asks about it.

no evidence and no free quantity: by equation (1.8) its temperature is fixed by the conductance, $T_2 = T_c + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$, a function of the state $\mathbf{z} = (T, g)^{\mathsf{T}}$, which keeps its two entries and its three rows. Chapter 4 carries $T_2$ instead as a third entry with a second physics row, the same formulation; with the same rows it gives the same answers. The set-point question asks for the same node with its supply air pinned at a set point $r$ not yet applied, $T_2(r) = r + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$, again a function of $g$ with no row of its own; the present condition, and with it the reading, stays at $T_c = 18°\mathrm{C}$.

**Which row acts on which entry.** The three rows act on the two entries of the state in a pattern fixed by the graph, before any number is computed. The physics row ties the temperature to the conductance and involves both; the sensor sees the temperature and not the conductance; the handbook speaks about the conductance and not the temperature; the scenario node adds no row. The Jacobian of the rows, the matrix of their derivatives with respect to the entries, can therefore be nonzero only where marked $*$:

$$\mathbf{J} = \begin{bmatrix} * & * \\ * & 0 \\ 0 & * \end{bmatrix} \begin{matrix} \text{physics row} \\ \text{reading row,} \\ \text{prior row} \end{matrix} \qquad \text{columns } z_1 = T,\ z_2 = g. \tag{1.9}$$

The zeros are structural: they hold at every state. Only the physics row has entries in both columns, and it is the one row through which a reading of the temperature reaches the conductance.

**The factor graph of the hall.** The pattern (1.9) is what a factor graph (Section 1.1.3) draws. Figure 1.4 draws it for the hall with its scenario node. The two solid circles are the entries of the state, $T$ and $g$. The black squares are the rows: the physics row of equation (1.5), linked to both entries; the reading row, linked to $T$ alone; and the prior row, linked to $g$ alone. The dashed circle is the full-load temperature $T_2$, joined to $g$ by a dashed line: it is a function of $g$ through equation (1.8), not an entry, and no row acts on it. The figure makes visible what the physical graph only implies: $T_2$ has no reading of its own and is connected to the evidence only through $g$, and $g$ is connected to the reading only through the physics row. The physical drawing has two supply-air boxes and two paths; the factor graph has no node for the supply air at all, since $T_c$ is pinned and appears only as a constant inside the physics row.

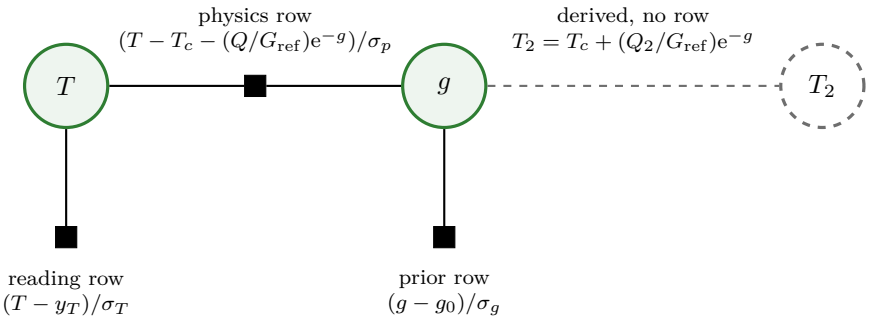



Figure 1.4: Factor graph of the equipment hall with its scenario node. Solid circles: the entries of the state, $T$ and $g$. Squares: the rows of the model, each linked to the entries it involves, the pattern (1.9) of the Jacobian. Dashed: the full-load temperature $T_2$, a function of $g$ with no row of its own, which reaches the evidence only through the shared $g$.

**The residual vector.** The state is the column vector of equation (1.3),

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} T \\ g \end{bmatrix},$$

with $T$ in °C and $g$ dimensionless. Three rows act on it, the physics row (1.5), the reading row (1.6) and the prior row, each divided by its standard deviation. Stacked into one residual vector, the three rows are

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} r_1(\mathbf{z}) \\ r_2(\mathbf{z}) \\ r_3(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} (T - T_c - (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g})/\sigma_p \\ (T - y_T)/\sigma_T \\ (g - g_0)/\sigma_g \end{bmatrix}, \tag{1.10}$$

with $T_c = 18°\mathrm{C}$, $Q = 30\,\mathrm{kW}$, $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$, the reading $y_T = 23.2°\mathrm{C}$ with $\sigma_T = 0.5\,\mathrm{K}$, and the handbook prior $g_0 = \ln(7.5/6) = 0.2231$ with $\sigma_g = 0.3$. Each entry of $\mathbf{r}$ is one row: a dimensionless number stating by how many standard deviations a candidate state violates one relation. The objective is half the sum of their squares,

$$\mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\,\mathbf{r}(\mathbf{z}) = \tfrac{1}{2}\sum_{k=1}^{3} r_k(\mathbf{z})^2, \tag{1.11}$$

which Chapter 3 shows to be minus the logarithm of the posterior up to a constant, so that the most probable state $\hat{\mathbf{z}}$, the maximum a posteriori (MAP) state, is the minimizer of $\mathcal{L}$. The computation takes $\sigma_p = 0.01\,\mathrm{K}$, fifty times tighter than the sensor, and Section 1.4.1 checks that the limit has been reached.

**The Jacobian.** The Jacobian $\mathbf{J}$ collects the derivative of every row with respect to every entry of the state: its entry in row $k$ and column $i$ is $\partial r_k/\partial z_i$, so it has one row per row of the model and one column per entry of the state, here three by two. Differentiating equation (1.10) entry by entry,

$$\mathbf{J}(\mathbf{z}) = \frac{\partial \mathbf{r}}{\partial \mathbf{z}} = \begin{bmatrix} 1/\sigma_p & (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g}/\sigma_p \\ 1/\sigma_T & 0 \\ 0 & 1/\sigma_g \end{bmatrix}, \tag{1.12}$$

the pattern (1.9) with its entries filled in. Each entry is a sensitivity in standard deviations: how far one row moves when one entry of the state moves by one unit. Only the entry in the first row and second column depends on the state; the others are constants. The first row is the

physics row's own sensitivities in its physical unit, rescaled. Differentiating the kilowatt form of equation (1.4),

$$\frac{\partial r_{\text{phys}}}{\partial T} = G_{\text{ref}}\,\mathrm{e}^{g} = G, \qquad \frac{\partial r_{\text{phys}}}{\partial g} = G_{\text{ref}}\,\mathrm{e}^{g}\,(T - T_c) = G\,(T - T_c),$$

in kW/K and kW; on the zero set the second equals the load $Q$ by equation (1.1). A one-kelvin error in the temperature unbalances the hall by $G$ kilowatts, and a unit change in $g$, a factor e in the conductance, by about $Q$ kilowatts to first order. Divided by $G$, to reach the temperature form, and by $\sigma_p$, the pair becomes $(1/\sigma_p,\ (Q/G)/\sigma_p)$, the first row of equation (1.12) on the zero set.

## 1.2 The posterior of the state (step 3)

All the questions of Table 1.2 are read from one posterior: the posterior of the state that the physics row, the reading row and the prior row of step 2 give, approximately the Laplace normal of $(T, g)$ and exactly a density along the physics curve. This section derives it and computes it for the hall; Section 1.3 reads every answer from it.

The objects of the fit recur in the specialized forms, and they are named here by their roles. The *MAP state* $\hat{\mathbf{z}} = (\hat{T}, \hat{g})^{\mathsf{T}}$, the maximum a posteriori state, is the minimum of the objective that the three rows put on the state, the most probable state; the *information matrix* $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ is the curvature that the rows give that objective at the minimum, with $\mathbf{J}$ the $3 \times 2$ Jacobian of the normalized rows; the *Laplace covariance* $\mathbf{\Sigma}$, a $2 \times 2$ matrix, is the inverse of that curvature after a term from the physics row's own curvature is added to it, and it makes the posterior of the state approximately normal, $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$; and a derived quantity $f$ is carried by its *gradient* $\mathbf{g} = \partial f/\partial \mathbf{z}$ at $\hat{\mathbf{z}}$ (bold, to tell it from the entry $g$), with the variance $\mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}$. Section 1.2.1 derives them from the rows. Beside them stands the exact posterior that this problem admits in the limit $\sigma_p \to 0$: on the physics curve the state has one free direction, and the posterior is the density of one entry, evaluated by quadrature. The Laplace route, a normal distribution centered on the MAP state, is what lets every answer be read from $\hat{\mathbf{z}}$ and $\mathbf{\Sigma}$ by operations that serve a plant of any size, and it misleads where the posterior is skewed along the curved ridge, which Section 1.4.2 measures against the exact posterior. The chapter reports the Laplace values, sets the exact value beside each answer with the gap, and collects the comparison in Section 1.4.

### 1.2.1 The posterior from the rows

The answers use the objects of one posterior: the MAP state $\hat{\mathbf{z}}$, the covariance $\mathbf{\Sigma}$ and the Jacobian $\mathbf{J}$ of the rows. The Jacobian is written out in Section 1.1.7; the other two are derived here for the hall. Chapter 3 derives these objects from the probability rules for Gaussian rows over any state, and Chapter 4 dissects the same objects in full on the hall with its scenario node carried as an entry, including the Laplace approximation and its curvature term. Here each is written out once for a reader meeting it for the first time.

**The posterior.** The ingredients are in hand: the unknown state $\mathbf{z} = (T, g)^{\mathsf{T}}$; the observation $y_T = 23.2°\text{C}$; the physical relation $Q = G\,(T - T_c)$; and the prior information $p(g)$. What is ultimately wanted is

$$p(\mathbf{z} \mid \mathbf{y}), \tag{1.13}$$

the *posterior distribution* of the state, conditional on the evidence $\mathbf{y}$, here the reading $y_T$ and the handbook value, and on the assumed physical model. It states which values of the state remain plausible after the physics, the reading and the prior information have been combined.

Chapter 3 derives Bayes' theorem, by which the posterior is proportional to the likelihood times the prior, $p(\mathbf{z} \mid \mathbf{y}) \propto p(y_T \mid \mathbf{z})\, p(\mathbf{z})$; here the reading row plays the part of the likelihood, and the physics row and the prior row on $g$ together play the part of the prior $p(\mathbf{z})$.

**The normal equations and the Gauss–Newton step.** Near a state $\mathbf{z}$, each row is replaced by its tangent, $\mathbf{r}(\mathbf{z}+\Delta\mathbf{z}) \approx \mathbf{r}(\mathbf{z}) + \mathbf{J}\,\Delta\mathbf{z}$, and the objective by $\frac{1}{2}\,\|\mathbf{r}+\mathbf{J}\,\Delta\mathbf{z}\|^2$, a quadratic in the step $\Delta\mathbf{z}$. Its gradient with respect to the step is $\mathbf{J}^\mathsf{T}(\mathbf{r}+\mathbf{J}\,\Delta\mathbf{z})$, and setting it to zero gives the *normal equations*

$$(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}, \tag{1.14}$$

a linear system of two equations in the two entries of the step. The Gauss–Newton method solves it, moves to $\mathbf{z}+\Delta\mathbf{z}$, evaluates $\mathbf{r}$ and $\mathbf{J}$ there, and repeats until the step no longer changes the state at the precision wanted. The vector $\mathbf{J}^\mathsf{T}\mathbf{r}$ is the gradient of the objective (1.11), so the iteration stops where that gradient vanishes, at the MAP state. The iteration starts here from the design state $\mathbf{z}^0 = (23^\circ\mathrm{C},\ 0)^\mathsf{T}$ of the forward calculation, at which the physics row is zero.

**The information matrix.** The matrix $\mathbf{J}^\mathsf{T}\mathbf{J}$ of the normal equations is the *information matrix* of the rows. It is a sum over the rows: writing $\mathbf{J}_k$ for the $k$-th row of $\mathbf{J}$,

$$\mathbf{J}^\mathsf{T}\mathbf{J} = \sum_{k=1}^{3} \mathbf{J}_k^\mathsf{T}\mathbf{J}_k,$$

so that each row contributes the outer product of its own sensitivities, and the precisions of independent rows add. For the hall the three outer products are of two kinds. The reading row and the prior row each have one nonzero entry, so each contributes one diagonal entry, $1/\sigma_T^2$ in the $T$–$T$ position and $1/\sigma_g^2$ in the $g$–$g$ position. The physics row has two, and its outer product $\mathbf{J}_1^\mathsf{T}\mathbf{J}_1$ fills the whole matrix, off-diagonal entries included: it is the only source of the coupling between the two entries, and it has rank one, with all its information along its own row and none along the tangent of the curve. The eigenvectors of $\mathbf{J}^\mathsf{T}\mathbf{J}$ give the directions in which the rows fix the state, and its eigenvalues how sharply: a large one across the curve, set by $1/\sigma_p^2$, and a small one along it, set by the reading and the prior, as the ridge below makes exact.

**The Laplace covariance and the curvature term.** The inverse of the information matrix is the covariance of the posterior: exactly so for rows linear in the state (Chapter 3), and as the *Laplace approximation*, the normal distribution fitted at the MAP state, for curved rows (Chapter 4). The Gauss–Newton matrix $\mathbf{J}^\mathsf{T}\mathbf{J}$ drops the curvature of the rows themselves. The full Hessian of the objective (1.11), its matrix of second derivatives, is

$$\frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}\,\partial \mathbf{z}^\mathsf{T}} = \mathbf{J}^\mathsf{T}\mathbf{J} + \sum_{k=1}^{3} r_k\, \frac{\partial^2 r_k}{\partial \mathbf{z}\,\partial \mathbf{z}^\mathsf{T}}, \tag{1.15}$$

and of the three rows only the physics row is curved, through $\mathrm{e}^{-g}$, with the single nonzero second derivative

$$\frac{\partial^2 r_1}{\partial g^2} = -\frac{(Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g}}{\sigma_p}.$$

The *curvature term* $r_1\,\partial^2 r_1/\partial g^2$ therefore adds to the $g$–$g$ entry alone. It does not vanish in the exact-physics limit, because $r_1$ shrinks in proportion to $\sigma_p$ while $\partial^2 r_1/\partial g^2$ grows as $1/\sigma_p$: at the MAP state the gradient $\mathbf{J}^\mathsf{T}\mathbf{r}$ is zero, its $T$ entry reads $r_1/\sigma_p + r_2/\sigma_T = 0$, so $r_1 = -(\sigma_p/\sigma_T)\,r_2$, and

$$r_1\,\frac{\partial^2 r_1}{\partial g^2} = \frac{r_2}{\sigma_T}\,\frac{Q}{G_{\mathrm{ref}}}\,\mathrm{e}^{-\hat{g}}$$

at every $\sigma_p$. It is the curvature of the physics curve itself, and it acts in the one direction, along the curve, in which the information is small. The *Laplace covariance* is the inverse of the full Hessian at the MAP state,

$$\boldsymbol{\Sigma} = \Big(\mathbf{J}^{\mathsf{T}}\mathbf{J} + \sum_k r_k \,\frac{\partial^2 r_k}{\partial \mathbf{z}\,\partial \mathbf{z}^{\mathsf{T}}}\Big)^{-1}_{\hat{\mathbf{z}}}, \qquad \mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma}); \tag{1.16}$$

a Newton iteration, which uses the full Hessian in place of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ in equation (1.14), reaches the same $\hat{\mathbf{z}}$.

**The ridge.** The spreads of the two entries are not independent of each other, and the reason is read off the physics row. Its Jacobian row, the derivative of equation (1.5) with respect to $(T, g)$, and the vector it annihilates are

$$\mathbf{J}_1 = \frac{1}{\sigma_p}\begin{bmatrix} 1 & (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g}\end{bmatrix}, \qquad \mathbf{n} = \begin{bmatrix}(Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g}\\ -1\end{bmatrix}, \qquad \mathbf{J}_1\,\mathbf{n} = 0. \tag{1.17}$$

A step along $\mathbf{n}$ leaves the physics row unchanged to first order, so $\mathbf{n}$ is the tangent to the zero set of the row, a change of $(Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$ kelvin of temperature per unit of $g$ with the opposite sign. The information the physics row contributes, $\mathbf{J}_1^{\mathsf{T}}\mathbf{J}_1$, is a matrix of rank one: it grows as $1/\sigma_p^2$ across the curve and is zero along $\mathbf{n}$. Along the curve only the reading row and the prior row inform the state, and by the pattern (1.9) each acts on one entry, with the contributions $\mathbf{J}_2^{\mathsf{T}}\mathbf{J}_2 = \mathrm{diag}(1/\sigma_T^2,\ 0)$ and $\mathbf{J}_3^{\mathsf{T}}\mathbf{J}_3 = \mathrm{diag}(0,\ 1/\sigma_g^2)$. With $n_1 = (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-\hat{g}}$ the first entry of $\mathbf{n}$ at the MAP state and $r_2$ the reading row there, the physics row's own curvature adds the term $\mathbf{C} = \mathrm{diag}(0,\ n_1 r_2/\sigma_T)$, derived above. With $\mathbf{u} = \mathbf{n}/\|\mathbf{n}\|$ the unit tangent, the information along the curve is

$$\lambda_{\parallel} = \mathbf{u}^{\mathsf{T}}(\mathbf{J}_2^{\mathsf{T}}\mathbf{J}_2 + \mathbf{J}_3^{\mathsf{T}}\mathbf{J}_3 + \mathbf{C})\mathbf{u} = \frac{n_1^2/\sigma_T^2 + 1/\sigma_g^2 + n_1 r_2/\sigma_T}{n_1^2 + 1}, \tag{1.18}$$

which does not depend on $\sigma_p$. The tangent $\mathbf{u}$ is orthogonal to $\mathbf{J}_1$, so in the orthonormal basis $(\mathbf{J}_1^{\mathsf{T}}/\|\mathbf{J}_1\|,\ \mathbf{u})$ the curvature of the objective has the entry $\lambda_{\parallel}$ along $\mathbf{u}$, an entry across the curve that grows as $1/\sigma_p^2$, and a coupling that stays bounded; its inverse, by the Schur complement of Chapter 3, leaves $1/\lambda_{\parallel}$ along $\mathbf{u}$ and nothing else as $\sigma_p \to 0$. The covariance therefore tends to the rank-one matrix $\boldsymbol{\Sigma} \to \mathbf{u}\,\mathbf{u}^{\mathsf{T}}/\lambda_{\parallel}$: the posterior collapses onto a segment of the tangent, a *ridge* of the objective. Then $\sqrt{\Sigma_{11}} \to |u_1|/\sqrt{\lambda_{\parallel}}$ and $\sqrt{\Sigma_{22}} \to |u_2|/\sqrt{\lambda_{\parallel}}$ are one spread along the curve seen through the two components of the tangent, and the correlation $\Sigma_{12}/\sqrt{\Sigma_{11}\Sigma_{22}}$ tends to the sign of $u_1 u_2$, which is $-1$. The two entries are not known separately: a hall slightly warmer than $\hat{T}$ is a hall with a slightly lower conductance, and the posterior allows the one only together with the other. The reading row has a zero in the column of $g$, and it narrows the conductance all the same, through the first component of $\mathbf{u}$.

> **Principle 1.2.** *Physical rows constrain the system; reading rows provide evidence about its unknown quantities. The posterior is where the two agree, weighted by the prior, and neither kind of row alone determines it. An exact physical row leaves the posterior only along its own zero set, and its signature in* $\boldsymbol{\Sigma}$ *is a correlation near* $-1$ *or* $+1$.

**The exact posterior along the curve.** In the limit $\sigma_p \to 0$ the posterior lies on the physics curve, and along the curve it can be computed without approximation. On the curve the temperature is a function of the conductance entry, $T(g) = T_c + (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$. The rows define the posterior relative to a base flat in $(T, g)$. Integrated across the curve, the physics row

contributes a factor that does not depend on $g$, because the row is written in kelvin and its slope in $T$, $1/\sigma_p$, does not depend on $g$; the base is therefore flat in $g$ along the curve, and what remains is the density of $g$ along the curve, the reading row and the prior row evaluated on it:

$$p(g \mid \mathbf{y}) \propto \exp\left[-\frac{1}{2}\left(\frac{T(g)-y_T}{\sigma_T}\right)^2-\frac{1}{2}\left(\frac{g-g_0}{\sigma_g}\right)^2\right]. \tag{1.19}$$

One-dimensional quadrature of this density over $g$ gives every probability and moment of a quantity derived from the state exactly, for $T$ through $T(g)$ and for $G$ through $G_{\mathrm{ref}}\,\mathrm{e}^{g}$.

**The answer read off.** An answer that is an entry of the state is read from $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ directly: the entry $\hat{z}_i$ of $\hat{\mathbf{z}}$, with the spread $\sqrt{\Sigma_{ii}}$. The relation between two entries is the correlation $\Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$, and the directions of the posterior are the eigenvectors of $\boldsymbol{\Sigma}$. A derived quantity $f$ is carried by its gradient, with the spread $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$, and a threshold on an entry is a normal tail of its marginal. Section 1.2.2 computes $\hat{\mathbf{z}}$, $\mathbf{J}$ and $\boldsymbol{\Sigma}$ for the hall.

### 1.2.2 The posterior of the hall

The hall's state estimation, the fit of the state to the three rows, has as inputs the pinned values $T_c = 18°\mathrm{C}$, $Q = Q_1 = 30\,\mathrm{kW}$ and $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$; the reading $y_T = 23.2°\mathrm{C}$ with $\sigma_T = 0.5\,\mathrm{K}$; the handbook prior $g_0 = 0.2231$ with $\sigma_g = 0.3$; and the tolerance $\sigma_p = 0.01\,\mathrm{K}$ of the physics row.

**The first step.** At the design state $\mathbf{z}^0 = (23°\mathrm{C},\ 0)^{\mathsf{T}}$ the Jacobian (1.12) is

$$\mathbf{J}(\mathbf{z}^0) = \begin{bmatrix} 100 & 500 \\ 2 & 0 \\ 0 & 3.333 \end{bmatrix}:$$

a one-kelvin change in $T$ moves the physics row by 100 and the reading row by 2, and a unit change in $g$ moves the physics row by 500 and the prior row by 3.33. The first row is the pair of kilowatt sensitivities $(G,\ Q) = (6,\ 30)$ divided by $G = 6\,\mathrm{kW/K}$ and by $\sigma_p$. The residual vector there is $\mathbf{r} = (0,\ -0.400,\ -0.744)$: the physics holds, the temperature lies 0.4 standard deviations below the reading, and $g$ lies 0.74 standard deviations below the handbook value. The two sides of the normal equations (1.14) are

$$\mathbf{J}(\mathbf{z}^0)^{\mathsf{T}}\mathbf{J}(\mathbf{z}^0) = \begin{bmatrix} 10004 & 50000 \\ 50000 & 250011 \end{bmatrix}, \qquad -\mathbf{J}(\mathbf{z}^0)^{\mathsf{T}}\mathbf{r}(\mathbf{z}^0) = \begin{bmatrix} 0.800 \\ 2.479 \end{bmatrix},$$

and the first step is their solution,

$$\Delta\mathbf{z} = \begin{bmatrix} 0.0684 \\ -0.0137 \end{bmatrix}.$$

**The iterations.** Table 1.4 lists the iterations. Each row is one iterate $\mathbf{z}^k$; the columns give its two entries, its three rows and the objective of equation (1.11).

The first step does almost all the work, because over a step this short the rows are nearly linear. It leaves the physics row slightly violated, $r_1 = -0.042$, through the curvature of $\mathrm{e}^{-g}$ that the tangent ignores; the second step restores it, and the third changes nothing at the printed precision. The MAP state is

$$\hat{\mathbf{z}} = \begin{bmatrix} \hat{T} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} 23.070°\mathrm{C} \\ -0.0139 \end{bmatrix}, \qquad \hat{G} = 6\,\mathrm{e}^{-0.0139} = 5.917\,\mathrm{kW/K}. \tag{1.20}$$

There the physics holds to $r_1\,\sigma_p = 5 \times 10^{-5}\,\mathrm{K}$, and the reading and the prior share the disagreement: the temperature lies 0.26 standard deviations below the reading and $g$ lies 0.79 standard deviations below the handbook value.

Table 1.4: Gauss–Newton iterations for the posterior of the equipment hall (equations (1.10) to (1.14)), with $\sigma_p = 0.01\,\mathrm{K}$, from the design state (23°C, 0).

| $k$ | $T$ [°C] | $g$ [–] | $r_1$ [–] | $r_2$ [–] | $r_3$ [–] | $\mathcal{L}$ [–] |
|---|---|---|---|---|---|---|
| 0 | 23.0000 | 0.0000 | 0.000 | −0.400 | −0.744 | 0.3566 |
| 1 | 23.0684 | −0.0137 | −0.042 | −0.263 | −0.789 | 0.3471 |
| 2 | 23.0701 | −0.0139 | 0.005 | −0.260 | −0.790 | 0.3460 |
| 3 | 23.0701 | −0.0139 | 0.005 | −0.260 | −0.790 | 0.3460 |

**The information matrix.** At the MAP state the Jacobian (1.12) is

$$\mathbf{J}(\hat{\mathbf{z}}) = \begin{bmatrix} 100 & 507.0 \\ 2 & 0 \\ 0 & 3.333 \end{bmatrix}, \qquad \frac{(Q/G_{\mathrm{ref}})\,\mathrm{e}^{-\hat{g}}}{\sigma_p} = \frac{5\,\mathrm{e}^{0.0139}}{0.01} = 507.0,$$

and the information matrix is the sum of the three outer products,

$$\mathbf{J}^{\mathsf{T}}\mathbf{J} = \underbrace{\begin{bmatrix} 10000 & 50701 \\ 50701 & 257056 \end{bmatrix}}_{\text{physics}} + \underbrace{\begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}}_{\text{reading}} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 11.11 \end{bmatrix}}_{\text{prior}} = \begin{bmatrix} 10004 & 50701 \\ 50701 & 257067 \end{bmatrix}.$$

The physics term has rank one: all its information lies along its own row (100, 507.0), across the physics curve, and none along the curve; only the reading's 4 and the prior's 11.11 inform along it. The eigenvalues of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ are 267067 across the curve and 4.27 along it, with the eigenvector $(0.9811,\ -0.1935)^{\mathsf{T}}$ of the small one pointing along the curve; the first two terms of equation (1.18) give it directly, $(5.070^2 \times 4 + 11.11)/(5.070^2 + 1) = 4.27$. The Gauss–Newton covariance is the inverse,

$$\mathbf{\Sigma}_{\mathrm{GN}} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1} = \begin{bmatrix} 0.2256 & -0.0445 \\ -0.0445 & 0.0088 \end{bmatrix},$$

with the standard deviations $\sqrt{0.2256} = 0.475\,\mathrm{K}$ and $\sqrt{0.0088} = 0.094$.

**The curvature term and the Laplace covariance.** The second derivative of the physics row at the MAP state is $\partial^2 r_1/\partial g^2 = -507.0$; multiplied by $r_1 = 0.0052$, it adds $-2.63$ to the $g$–$g$ entry, $257067 \to 257064$, the same value as $(r_2/\sigma_T)(Q/G_{\mathrm{ref}})\,\mathrm{e}^{-\hat{g}} = (-0.2598/0.5) \times 5.0701$. Projected on the along-curve eigenvector $(0.9811,\ -0.1935)^{\mathsf{T}}$, the term is $-2.63 \times 0.1935^2 = -0.099$ against the along-curve information 4.27, a loss of 2.3 % in the one direction in which the information is small; it raises the variances by 2.4 % and the standard deviations by 1.2 %. The Laplace covariance (1.16) is

$$\mathbf{\Sigma} = \begin{bmatrix} 0.23096 & -0.045552 \\ -0.045552 & 0.0089881 \end{bmatrix}, \qquad \sqrt{\Sigma_{11}} = 0.4806\,\mathrm{K}, \quad \sqrt{\Sigma_{22}} = 0.09481. \tag{1.21}$$

These are the Laplace values used in the rest of the chapter, quoted as 0.481 K and 0.0948; Gauss–Newton alone gives 0.475 K and 0.094. The correlation of the two entries, computed from the unrounded $\mathbf{\Sigma}$, is

$$\frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\,\Sigma_{22}}} = \frac{-0.045552}{0.48058 \times 0.094806} = -0.9998.$$

**The ridge.** The physics row of $\mathbf{J}(\hat{\mathbf{z}})$, $(100,\ 507.0)$, is annihilated by $\mathbf{n} = (5.070,\ -1)^{\mathsf{T}}$, equation (1.17) at $\hat{g}$, and the long axis of $\boldsymbol{\Sigma}$ is $\mathbf{u} = (0.9811,\ -0.1935)^{\mathsf{T}}$, the same direction normalized. By equation (1.18),

$$\lambda_{\parallel} = \frac{5.070^2 \times 4 + 11.11 + 5.070 \times (-0.260)/0.5}{5.070^2 + 1} = 4.168,$$

$$\frac{0.9811}{\sqrt{4.168}} = 0.481\,\mathrm{K}, \quad \frac{0.1935}{\sqrt{4.168}} = 0.0948,$$

the square roots of the two diagonal entries of equation (1.21).

Figure 1.5 shows the geometry. The horizontal axis is the log-conductance $g$ and the vertical axis the hall temperature $T$, the two entries of the state, so that every point of the plane is a candidate state. The green curve is the zero set of the physics row: the states the physics allows. The gray band is the set of temperatures within one standard deviation of the reading, $23.2 \pm 0.5°\mathrm{C}$, and the dashed line inside it the reading itself. Points are labeled by their plane coordinates $(g, T)$, horizontal first. The three black dots are the design state, $(g, T) = (0, 23)$; the state on the curve that matches the reading exactly, $(g, T) = (-0.039, 23.2)$; and, between them, the MAP state of equation (1.20), $(g, T) = (-0.014, 23.07)$. The purple segment is the one-standard-deviation ellipse of the covariance of equation (1.21), centered on the MAP state; it is drawn as a thick line because its half-width across the curve, about 0.002 in $g$, is below the resolution of the figure.

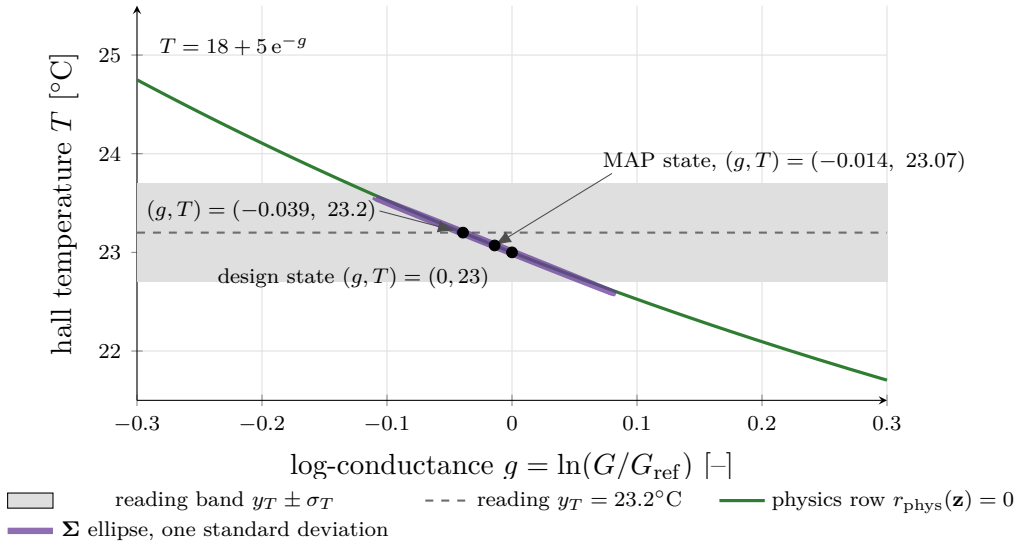



Figure 1.5: The state plane of the equipment hall with $T_c = 18°\mathrm{C}$, $Q = 30\,\mathrm{kW}$ and $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$. Green: the zero set of the physics row, $T = 18 + 5\,\mathrm{e}^{-g}$ (equation (1.4)). Gray: the reading $y_T = 23.2°\mathrm{C}$ with its band of one standard deviation, $\sigma_T = 0.5\,\mathrm{K}$ (equation (1.6)). The plausible states lie on the arc of the curve inside the band. Purple: the one-standard-deviation ellipse of the Laplace covariance (equation (1.21)) around the MAP state, $(g, T) = (-0.014,\ 23.07)$, collapsed onto the curve by the physics row. Points are labeled $(g, T)$, horizontal coordinate first.

The plausible states lie on the arc of the curve inside the band, from $g = -0.131$ to $g = 0.062$ ($G = 5.26$ to $6.38\,\mathrm{kW/K}$), and the MAP state sits on it between the reading-match point and the design state, pulled toward the handbook value $g_0 = 0.2231$ by the prior row.

The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of equations (1.20) and (1.21) is what step 4 uses: every answer of the chapter is computed from it.

## 1.3 The answers (step 4)

The first five questions of Table 1.2 take two general forms, written in the book's notation: $p(\mathbf{z} \mid \mathbf{y})$ is the posterior density of the state, $f(\mathbf{z})$ a quantity derived from the state and $f^*$ a threshold on it. The added-sensor question, a design question about a row not yet attached, is derived in Section 1.6.1. Each is answered from the posterior of Section 1.2: the engineer's question, the general form of its category, the form it takes for the hall, the quantities of the posterior it uses, and the hall's numbers. The questions name the limit of 24°C at the present load and the limit of 27.5°C at $Q_2 = 60\,\mathrm{kW}$.

### 1.3.1 The hall temperature and the process gain

**Control engineer's questions.** What temperature is the hall running at now? What is the process gain to the load? The first asks for the temperature of the hall at the present load, given its one reading and the handbook value; the second for the kelvin the hall gains per kilowatt of load, which rests on the conductance of its air path, the quantity the engineer compares with the design value.

**General form.** An estimation question about a quantity $f(\mathbf{z})$ derived from the state asks for its posterior distribution, the probability that the quantity lies at or below each value $f^*$ (the value at which it equals one half is the median),

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \tag{1.22}$$

where $\mathbf{1}[\cdot]$ is one when the event holds and zero otherwise, summarized by a location and a spread,

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad \mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y}) = \mathbb{E}\big[(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}])^2 \mid \mathbf{y}\big]. \tag{1.23}$$

An entry of the state is the case $f(\mathbf{z}) = z_i$. These are the estimation forms (2.9) and (2.10) that Chapter 2 develops.

**Specialized form.** The two questions are equations (1.22) and (1.23), the forms (2.9) and (2.10), with the state $\mathbf{z} = (T, g)^\mathsf{T}$ of the hall, the derived quantity $f(\mathbf{z}) = z_1 = T$ for the temperature and $f(\mathbf{z}) = K_Q = \mathrm{e}^{-z_2}/G_{\mathrm{ref}}$ for the process gain, through $f(\mathbf{z}) = z_2 = g$ and $f(\mathbf{z}) = G = G_{\mathrm{ref}}\,\mathrm{e}^{z_2}$, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of equations (1.20) and (1.21).

*The temperature.* Here $f(\mathbf{z}) = z_1 = T$, the first entry of the state. Under the Laplace posterior the marginal of an entry of a normal vector is normal, with the entry's mean and the diagonal entry of the covariance, so

$$T \mid \mathbf{y} \approx \mathcal{N}(\hat{z}_1, \Sigma_{11}),$$

and the answer is read off without any transformation: the first entry of $\hat{\mathbf{z}}$, with the spread $\sqrt{\Sigma_{11}}$.

*The conductance and the process gain.* The second entry $z_2 = g$ is read off in the same way, $g \mid \mathbf{y} \approx \mathcal{N}(\hat{z}_2, \Sigma_{22})$. The conductance itself, $G = G_{\mathrm{ref}}\,\mathrm{e}^{z_2}$, is a derived quantity of the state. Linearized at the MAP state, its variance is $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$ with the gradient

$$\mathbf{g} = \frac{\partial G}{\partial \mathbf{z}} = \begin{bmatrix} \partial G/\partial T \\ \partial G/\partial g \end{bmatrix}_{\hat{\mathbf{z}}} = \begin{bmatrix} 0 \\ G_{\mathrm{ref}}\,\mathrm{e}^{\hat{g}} \end{bmatrix} = \begin{bmatrix} 0 \\ \hat{G} \end{bmatrix}, \qquad \sigma_G^2 = \mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} = \hat{G}^2\,\Sigma_{22}, \tag{1.24}$$

the delta-method spread, written $\sigma_G$. The ratio $\sigma_G/\hat{G} = \sqrt{\Sigma_{22}}$: the spread of the log-conductance is the fractional spread of the conductance, as Section 1.1.4 anticipated. The process gain to the load is the derivative of the forward relation (1.2),

$$K_Q = \frac{\partial T}{\partial Q} = \frac{1}{G} = \frac{\mathrm{e}^{-z_2}}{G_{\mathrm{ref}}}, \qquad \mathbf{g}_K = \begin{bmatrix} 0 \\ -\hat{K}_Q \end{bmatrix}, \qquad \sigma_K^2 = \mathbf{g}_K^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}_K = \hat{K}_Q^2 \, \Sigma_{22}, \tag{1.25}$$

a second derived quantity of the same entry, with the same fractional spread as the conductance and the opposite sign in its gradient: a larger conductance is a smaller gain.

The value beside each comes from the exact posterior: the moments (1.23) under the density (1.19), evaluated by one-dimensional quadrature over $g$, with $T$ through $T(g) = T_c + (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$, $G$ through $G_{\mathrm{ref}}\,\mathrm{e}^{g}$ and $K_Q$ through $\mathrm{e}^{-g}/G_{\mathrm{ref}}$, and the median of each, the value $f^*$ at which equation (1.22) equals one half. The Laplace location, the entry of $\hat{\mathbf{z}}$, is the median of the Laplace normal, and it is set beside the exact median, like with like; the exact mode of $g$ is $\hat{g}$ itself, and the exact mean is given as well, labeled as a mean. They differ by the skew of the posterior along the curve.

**Building blocks.** The MAP state $\hat{\mathbf{z}} = (23.070\,^{\circ}\mathrm{C},\ -0.0139)^{\mathsf{T}}$ and the Laplace covariance $\boldsymbol{\Sigma}$, curvature term included, with $\sqrt{\Sigma_{11}} = 0.4806\,\mathrm{K}$ and $\sqrt{\Sigma_{22}} = 0.09481$ (equations (1.20) and (1.21)); the gradients (1.24) and (1.25); and, for the exact values beside them, the density (1.19) and a one-dimensional quadrature over $g$.

**The hall temperature now.** The hall temperature is the first entry of the state: $\hat{T} = 23.070\,^{\circ}\mathrm{C}$ with $\sqrt{\Sigma_{11}} = 0.481\,\mathrm{K}$ (Gauss–Newton alone, $0.475\,\mathrm{K}$). The value beside it comes from the exact posterior: the median of $T$ along the curve is $23.028\,^{\circ}\mathrm{C}$, $0.042\,\mathrm{K}$ below $\hat{T}$ (the mean is $23.030\,^{\circ}\mathrm{C}$), with the standard deviation $0.4813\,\mathrm{K}$, $0.0007\,\mathrm{K}$ wider.

**The effective conductance.** The conductance entry is $\hat{g} = -0.0139$ with $\sqrt{\Sigma_{22}} = 0.0948$. By equation (1.24), with $\mathbf{g} = (0,\ 5.917)^{\mathsf{T}}$,

$$\hat{G} = 5.917\,\mathrm{kW/K}, \qquad \sigma_G = \sqrt{\mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}} = 5.917\,\sqrt{0.0089881} = 0.561\,\mathrm{kW/K},$$

a band of 9.5 % of its value, against a factor $\mathrm{e}^{0.3} = 1.35$ in the handbook before the reading. The values beside them come from the exact posterior: the conductance entry has the median $-0.0056$, $0.0083$ above $\hat{g}$ (the mean is $-0.0013$), and the standard deviation $0.0967$, $0.0019$ wider; the conductance has the median $5.967\,\mathrm{kW/K}$, $0.049\,\mathrm{kW/K}$ above $\hat{G}$ (the mean is $6.020\,\mathrm{kW/K}$), and the standard deviation $0.591\,\mathrm{kW/K}$, $0.030\,\mathrm{kW/K}$ or 5 % wider. The exact posterior extends further toward large conductances than the normal at $\hat{\mathbf{z}}$ does, the skew that the full-load question below meets in its tail.

**The process gain.** By equation (1.25), with $\mathbf{g}_K = (0,\ -0.1690)^{\mathsf{T}}$,

$$\hat{K}_Q = \frac{1}{5.917} = 0.1690\,\mathrm{K/kW}, \qquad \sigma_K = \sqrt{\mathbf{g}_K^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}_K} = 0.1690\,\sqrt{0.0089881} = 0.0160\,\mathrm{K/kW},$$

against the $0.1733\,\mathrm{K/kW}$ of the commissioning sheet, which takes the reading as exact. The values beside them come from the exact posterior: the median is $0.1676\,\mathrm{K/kW}$ (the mean is $0.1677\,\mathrm{K/kW}$), with the standard deviation $0.0160\,\mathrm{K/kW}$.

### 1.3.2 The limits at the present load and at 60 kilowatts

**Control engineer's questions.** Is the hall within its 24°C limit now? Can the hall carry 60 kW within 27.5°C at this set point? The first asks about the hall as it runs now. The second is the owner's request, which for the engineer reads

*What is the probability that the equipment hall will remain below* 27.5°C *when the load increases to* 60 kW*?*

Its unknown is the temperature $T_2$ of the scenario node at the full load of 60 kW, and over the state of the graph it is

$$\Pr(T_2 < 27.5^\circ\mathrm{C} \mid \mathbf{y}), \tag{1.26}$$

with the unknown, the full-load temperature, on the left of the bar and the evidence, the reading of the present condition and the handbook value, on the right; the graph and its pinned inputs are conditions of the model throughout.

**General form.** A prediction under a threshold is a tail probability of a derived quantity under the posterior, the posterior mass of the states that satisfy the event,

$$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) < f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}. \tag{1.27}$$

This is the threshold form (2.12) of Chapter 2, written for an upper limit.

**Specialized form.** The two questions are equation (1.27), the form (2.12), with the state $\mathbf{z} = (T, g)^{\mathsf{T}}$ of the hall, $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of equations (1.20) and (1.21), and the derived quantity $f(\mathbf{z})$ and the threshold $f^*$ given for each below: $f(\mathbf{z}) = T$ with $f^* = 24^\circ\mathrm{C}$, and $f(\mathbf{z}) = T_2$ with $f^* = T_{\max} = 27.5^\circ\mathrm{C}$. Two choices enter here. The full-load condition is taken to share the conductance of the present condition, with the load and the supply air pinned at both: this is what makes a condition the hall has never run predictable from a reading at 30 kW, and it fails if the conductance changes with the load, for instance through an airflow that rises with it (Section 1.1.4), or if the valve reaches its stop and the loop can no longer hold $T_c$ at 60 kW (Section 1.6.5), and no check of this chapter can test either because nothing has been observed at 60 kW. The full-load answer is the steady temperature at 60 kW: this is what lets the event be written on the conductance alone, and it says nothing about the period just after the load rises, while the hall's stored heat still matters (Section 1.6.5).

*The present load.* Here $f(\mathbf{z}) = z_1 = T$ and $f^* = 24^\circ\mathrm{C}$. Under the Laplace posterior the tail is a normal tail of the first entry,

$$\Pr(T < 24^\circ\mathrm{C} \mid \mathbf{y}) \approx \Phi\!\left(\frac{24 - \hat{z}_1}{\sqrt{\Sigma_{11}}}\right), \tag{1.28}$$

where $\Phi$ is the standard normal cumulative distribution function. The value beside it comes from the exact posterior: on the curve the event $T < 24^\circ\mathrm{C}$ is the event $T(g) < 24^\circ\mathrm{C}$, that is $g > -\ln\big((24 - T_c)\, G_{\mathrm{ref}}/Q\big) = -\ln 1.2$, and its probability is the integral of equation (1.19) above that value.

*The full load.* Here $f(\mathbf{z}) = T_2$, the temperature of the scenario node. By equation (1.8) it is a function of the second entry alone, $T_2 = T_c + (Q_2/G_{\mathrm{ref}})\, \mathrm{e}^{-z_2}$, and on the zero set of the physics row it is also a function of the first: dividing equation (1.8) by equation (1.7) removes the conductance,

$$\frac{T_2 - T_c}{T - T_c} = \frac{Q_2}{Q_1} \qquad \Longrightarrow \qquad T_2 = T_c + \frac{Q_2}{Q_1}\,(T - T_c). \tag{1.29}$$

At twice the load the excess over the supply air is twice as large, whatever the conductance: a present temperature of 23°C predicts $T_2 = 18 + 2 \times 5 = 28$°C, and a present temperature of 23.2°C predicts 28.4°C. Readings taken during the first operating condition therefore inform the second, although the hall has never been operated at full load; this is how a plant model links the telemetry of today to predictions under operating conditions that have not yet occurred. In the vector form the link is the tangent $\mathbf{n}$ of equation (1.17): a step along the curve moves $T$ by $(Q_1/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$ and $T_2$ by $(Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$ per unit of $-g$, in the ratio $Q_2/Q_1$.

The event itself needs no posterior to be rewritten. The full-load temperature $T_2 = T_c + Q_2/G$ decreases in $G$, so for a limit $T_{\max} > T_c$ the event $T_2 < T_{\max}$ is, in turn, an event on the conductance, on the conductance entry and, by equation (1.29), on the present temperature:

$$T_2 < T_{\max} \iff G > G^* \iff g > g^* = \ln \frac{G^*}{G_{\mathrm{ref}}} \iff T < T^*, \tag{1.30}$$

with the thresholds

$$G^* = \frac{Q_2}{T_{\max} - T_c}, \tag{1.31}$$

$$T^* = T_c + \frac{Q_1}{Q_2}\,(T_{\max} - T_c). \tag{1.32}$$

For $T_{\max} \le T_c$ the probability is zero, since $T_2 > T_c$ for every positive conductance. The inequalities of equation (1.30) describe the same set of states on the physics curve, so their probabilities are equal under any distribution on the curve; the equivalence holds under the assumed steady-state model, with known load and supply-air temperature. The question about a condition that has never been operated has become a question about the unknown conductance, an entry of the state that the present condition informs, before any data are processed.

Figure 1.6 draws the equivalence (1.30) on the forward relation (1.2). The horizontal axis is the conductance $G$ in kW/K and the vertical axis the hall temperature in °C. The solid blue curve is the hall at the load of 30 kW, the dashed green curve the same hall at 60 kW; the dot on the blue curve is the forward calculation of Section 1.1.4, $T = 23$°C at $G = 6$ kW/K, and the dot above it the same conductance at 60 kW, 28°C. The red line is the limit of 27.5°C, and the dotted vertical line marks the threshold $G^*$ of equation (1.31), where the dashed curve crosses the limit. The dashed green curve lies below the red line exactly to the right of $G^*$: because the curve is strictly decreasing, the event "below the limit" on the vertical axis is the event "above $G^*$" on the horizontal axis.

Under the Laplace posterior the probability is a normal tail of the second entry, the entry onto which the physics maps the event,

$$\Pr(T_2 < T_{\max} \mid \mathbf{y}) = \Pr(g > g^* \mid \mathbf{y}) \approx \Phi\!\left(\frac{\hat{z}_2 - g^*}{\sqrt{\Sigma_{22}}}\right). \tag{1.33}$$

The same event written on the first entry, $T < T^*$, has the same exact probability, but under the Laplace normal, whose ridge is the straight tangent rather than the curve, its tail differs slightly; the chapter reports the tail on $g$, the entry on which the full-load temperature depends alone. The value beside it comes from the exact posterior: the probability is the integral of equation (1.19) above $g^*$.

**Principle 1.3.** *When a quantity in a question is a monotone function of one entry of the state, a threshold on the quantity is a threshold on that entry, and the two events have the same probability under every posterior. The transformation uses the physics only, and it can be made before any data are processed.*

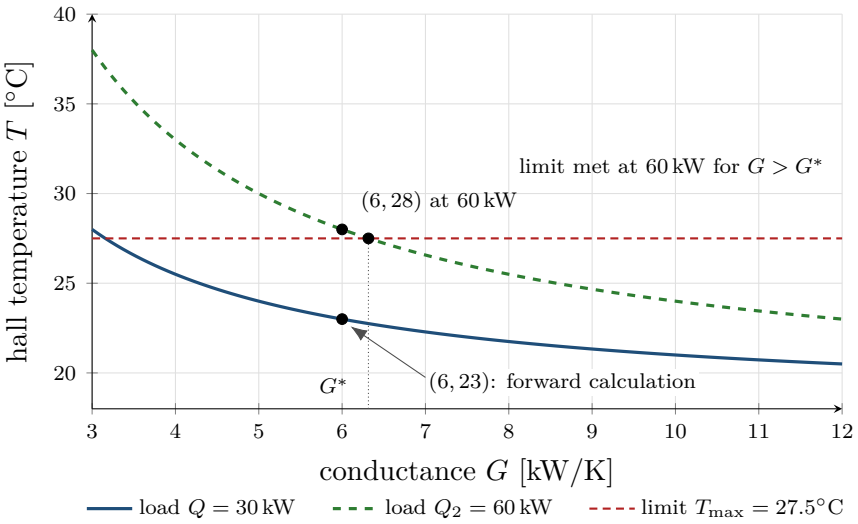



Figure 1.6: The forward relation $T = T_c + Q/G$ (equation (1.2)) with $T_c = 18$°C, for the loads $Q = 30$ kW and $Q_2 = 60$ kW. Dots: the forward calculation $T = 23$°C at $G = 6$ kW/K; the same conductance at 60 kW, 28°C; and the threshold $G^* = Q_2/(T_{\max} - T_c)$ of equation (1.31) at which the 60 kW hall reaches the limit of 27.5°C.

**Building blocks.** The entries $\hat{z}_1 = 23.070$°C and $\hat{z}_2 = -0.0139$ of the MAP state and the standard deviations $\sqrt{\Sigma_{11}} = 0.481$ K and $\sqrt{\Sigma_{22}} = 0.0948$ of the covariance (equations (1.20) and (1.21)); the threshold of equation (1.31) for the requested $T_{\max} = 27.5$°C and $Q_2 = 60$ kW, and for other limits and loads; the normal cumulative distribution $\Phi$; and, for the exact values beside them, the density (1.19) integrated over $g$ above each threshold.

**The hall below 24 degrees.** By equation (1.28),

$$\Pr(T < 24\text{°C} \mid \mathbf{y}) \approx \Phi\left(\frac{24 - 23.070}{0.481}\right) = \Phi(1.93) = 0.973.$$

The value beside it comes from the exact posterior: the mass of equation (1.19) above $g = -\ln 1.2 = -0.1823$ is 0.977, 0.004 higher.

**The hall at 60 kilowatts.** For the requested limit $T_{\max} = 27.5$°C at $Q_2 = 60$ kW, equations (1.31) and (1.32) give

$$G^* = \frac{60}{27.5 - 18} = \frac{60}{9.5} = 6.316\,\text{kW/K}, \qquad g^* = \ln\frac{6.316}{6} = 0.0513,$$

$$T^* = 18 + \tfrac{1}{2} \times 9.5 = 22.75\text{°C}:$$

the temperature at 60 kW stays below 27.5°C if and only if the temperature at 30 kW is below 22.75°C; on the curve of Figure 1.5, the states with $g > g^*$ are those below 22.75°C. By equation (1.33),

$$\Pr(T_2 < 27.5\text{°C} \mid \mathbf{y}) \approx \Phi\left(\frac{-0.0139 - 0.0513}{0.0948}\right) = \Phi(-0.688) = 0.246,$$

with the MAP conductance 5.917 kW/K to the left of $G^*$ in Figure 1.6. The value beside it comes from the exact posterior: the mass of equation (1.19) above $g^* = 0.0513$ is 0.281, 0.036 higher. The event lies in the tail of $g$ toward large conductances, the side on which the exact posterior extends further than the normal at $\hat{\mathbf{z}}$, so the state route understates it. The gap does

not change the decision: on both routes the hall meets the limit at 60 kW with probability of about one in four.

Table 1.5 evaluates the same blocks for other limits and loads. Each row is a limit $T_{\max}$ and a future load $Q_2$; the columns give the threshold conductance $G^*$, the threshold log-conductance $g^*$, the threshold on the present temperature $T^*$, the probability of equation (1.33), which the chapter reports, and beside it the exact value, the mass of equation (1.19) above $g^*$. One row of the arithmetic, for $T_{\max} = 27°\mathrm{C}$ and $Q_2 = 60\,\mathrm{kW}$: $G^* = 60/9 = 6.667\,\mathrm{kW/K}$, $g^* = \ln(6.667/6) = 0.1054$, $T^* = 18 + 0.5 \times 9 = 22.5°\mathrm{C}$, and $\Phi((-0.0139 - 0.1054)/0.0948) = \Phi(-1.258) = 0.104$.

Table 1.5: Equivalent thresholds and the full-load probability, with $T_c = 18°\mathrm{C}$, $Q_1 = 30\,\mathrm{kW}$ and $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$ (equations (1.31) and (1.32)); the probability by the Laplace tail (1.33), with $\hat{g} = -0.0139$ and $\sqrt{\Sigma_{22}} = 0.0948$, and exactly, by quadrature of equation (1.19). The row 27.5°C at 60 kW is the owner's request; the others are added to show the trend.

| Limit $T_{\max}$ [°C] | Load $Q_2$ [kW] | $G^*$ [kW/K] | $g^*$ [–] | $T^*$ [°C] | Laplace [–] | Exact [–] |
|---|---|---|---|---|---|---|
| 26 | 60 | 7.500 | 0.2231 | 22.00 | 0.0062 | 0.0155 |
| 27 | 60 | 6.667 | 0.1054 | 22.50 | 0.1042 | 0.1354 |
| 27.5 | 60 | 6.316 | 0.0513 | 22.75 | 0.2458 | 0.2815 |
| 29 | 60 | 5.455 | −0.0953 | 23.50 | 0.8047 | 0.8355 |
| 27.5 | 45 | 4.737 | −0.2364 | 24.33 | 0.9905 | 0.9964 |
| 27.5 | 75 | 7.895 | 0.2744 | 21.80 | 0.0012 | 0.0049 |

A stricter limit, or a larger future load, demands a larger conductance and a cooler present temperature; the first five columns follow from the physics alone. The Laplace tail is below the exact one in every row, and the gap grows in ratio toward the far tail: a factor of 2.5 to 4 in the two smallest rows.

### 1.3.3 The set point at 60 kilowatts

**Control engineer's question.** How far must the set point move to hold the hall within 27.5°C at 60 kW? The commissioning sheet answers 17.1°C, with no probability attached.

**General form.** The question is a threshold on a derived quantity under the posterior, the tail (1.27), read at a scenario node whose pinned input $r$ is the value being asked about,

$$\Pr(f(\mathbf{z}; r) < f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}; r) < f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}. \tag{1.34}$$

**Specialized form.** The question is equation (1.34), the form (2.12), with $f(\mathbf{z}; r) = T_2(r)$, $f^* = T_{\max} = 27.5°\mathrm{C}$ and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3. The scenario node is closed loop: it pins the set point $r$ and leaves the valve to the loop, and the reading stays at the present set point, so that $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ are those of equations (1.20) and (1.21) for every $r$. The controller is taken to hold $r$ exactly: this is what lets the set point be read as a number with no spread, and it fails at the valve's stop, and if the discharge sensor is off, since the loop holds its reading and not the air (Chapter 6). The derived quantity and its two gains are

$$T_2(r) = r + \frac{Q_2}{G_{\mathrm{ref}}}\,\mathrm{e}^{-z_2}, \qquad \frac{\partial T_2}{\partial r} = 1, \qquad \frac{\partial T_2}{\partial \mathbf{z}} = \begin{bmatrix} 0 \\ -(Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-z_2} \end{bmatrix}. \tag{1.35}$$

The set-point gain is exactly 1 K/K for every $g$, because $T_2(r)$ is $r$ plus a term that does not contain $r$, the conductance being taken as independent of the supply air. Below the floor, the set point at which the valve reaches its stop (Section 1.6.5), the gain is zero. The gradient on

the state does not contain $r$. With the sensor true, a pinned set point therefore adds no spread and moves only the threshold: as in equation (1.30), the event is an event on the conductance entry,

$$T_2(r) < T_{\max} \iff g > g^*(r) = \ln \frac{Q_2}{(T_{\max} - r)\, G_{\text{ref}}}, \qquad \Pr(T_2(r) < T_{\max} \mid \mathbf{y}) \approx \Phi\!\left( \frac{\hat{z}_2 - g^*(r)}{\sqrt{\Sigma_{22}}} \right), \tag{1.36}$$

the tail (1.33) with $g^*$ moved. The value beside it comes from the exact posterior: the mass of equation (1.19) above $g^*(r)$. This is a threshold question read at a pinned input; Chapter 6 derives the chance constraint and its back-off; Chapter 16 what a tighter posterior is worth.

**Building blocks.** The entry $\hat{z}_2 = -0.0139$ and the spread $\sqrt{\Sigma_{22}} = 0.0948$ (equations (1.20) and (1.21)); the threshold $g^*(r)$ of equation (1.36); the normal cumulative distribution $\Phi$; and, for the exact values, the density (1.19) integrated above $g^*(r)$.

**The set points of the hall.** Table 1.6 evaluates equation (1.36). Each row is one set point; the columns give the threshold $g^*(r)$, the probability by the state route, which the chapter reports, and the exact value beside it. The first row is the present set point, the probability of Section 1.3.2. One row of the arithmetic, the sheet's set point: $g^*(17.10) = \ln\big(60/((27.5 - 17.10) \times 6)\big) = -0.0392$, the log-conductance of the reading taken as exact, and $\Phi((-0.0139 + 0.0392)/0.0948) = 0.605$.

Table 1.6: The set point at 60 kW within 27.5°C, from equation (1.36), with $\hat{g} = -0.0139$ and $\sqrt{\Sigma_{22}} = 0.0948$, and exactly, by quadrature of equation (1.19). The row 17.10°C is the commissioning sheet's set point. The table reads the question at a pinned supply air; below the coil's floor of 17.75°C (Section 1.6.5) the supply air stays at the floor.

| Set point $r$ [°C] | $g^*(r)$ [–] | State route [–] | Exact [–] |
|---:|---:|---:|---:|
| 18.00 | 0.0513 | 0.2458 | 0.2815 |
| 17.75 | 0.0253 | 0.3395 | 0.3752 |
| 17.50 | 0.0000 | 0.4416 | 0.4768 |
| 17.10 | −0.0392 | 0.6052 | 0.6393 |
| 17.00 | −0.0488 | 0.6435 | 0.6772 |
| 16.50 | −0.0953 | 0.8047 | 0.8355 |
| 16.00 | −0.1398 | 0.9078 | 0.9320 |
| 15.50 | −0.1823 | 0.9622 | 0.9775 |

At a pinned supply air, the sheet's 17.1°C holds the hall with probability 0.605, or 0.639 exactly: it is the set point at which the limit is met if the conductance is the one the reading alone implies, and the posterior places the rest of its mass on the wrong side of that value. The largest set point at which the probability reaches 0.95 is 15.65°C by the state route, 2.35 K below the present one, and 15.85°C exactly, a gap of 0.20 K. The state route errs on the cold side here, by the same skew that makes its tail low in Table 1.5. Whether the coil can deliver any of these set points is outside the graph of this chapter; Section 1.6.5 states the limit.

## 1.4 The checks (step 5)

The fit rests on two approximations and one assumption: a physics row with a finite tolerance in place of an exact law, a normal distribution in place of the posterior along a curved ridge,

and a reading and a handbook value that are compatible with each other. Each is set against a full calculation, and the weight of the handbook value is measured by removing it.

### 1.4.1 The exact-physics limit

Table 1.7 repeats the computation at three model tolerances. Each row is one $\sigma_p$; the columns are the two entries of $\hat{\mathbf{z}}$, their Laplace standard deviations and their correlation. From 0.1 to 0.01 K the conductance entry moves by 0.001, its spread by 0.002 and the correlation from $-0.979$ to $-0.9998$; from 0.01 to 0.001 K nothing moves at the printed precision but the last digit of the correlation. The limit is reached at $\sigma_p = 0.01\,\mathrm{K}$, and the correlation tends to $-1$ as equation (1.18) requires.

Table 1.7: The posterior of the equipment hall at three model tolerances $\sigma_p$ of the physics row: MAP state, Laplace standard deviations (with the curvature term) and correlation. The exact-physics limit $\sigma_p \to 0$ is reached at 0.01 K.

| $\sigma_p$ [K] | $\hat{T}$ [°C] | $\hat{g}$ [–] | $\sqrt{\Sigma_{11}}$ [K] | $\sqrt{\Sigma_{22}}$ [–] | correlation [–] |
|---|---|---|---|---|---|
| 0.1 | 23.0705 | −0.0130 | 0.4806 | 0.0966 | −0.9790 |
| 0.01 | 23.0701 | −0.0139 | 0.4806 | 0.0948 | −0.9998 |
| 0.001 | 23.0701 | −0.0139 | 0.4806 | 0.0948 | −1.0000 |

### 1.4.2 The Laplace posterior against the exact posterior

In the limit $\sigma_p \to 0$ the exact posterior is the density (1.19) along the physics curve, and one-dimensional quadrature of it gives every quantity of the questions; the event $T < 24$°C is the event $g > -\ln 1.2 = -0.1823$ on the curve, and the full-load event is $g > g^* = 0.0513$. Table 1.8 collects the pairs that step 4 computed, the Laplace value it reports beside the exact one. Each row is one quantity of the posterior, with its unit; the third column is its Laplace value, read from $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ of equations (1.20) and (1.21), and the fourth its exact value from the quadrature. The location of each quantity is its median in both columns, the Laplace median being the value at $\hat{\mathbf{z}}$; the exact mean has a row of its own.

For the temperature the approximation is adequate: the tail probability differs by 0.004 and the standard deviation by 0.0007 K. The posterior is slightly skewed along the curve, so its mode, median and mean differ: the mode of $g$ is $\hat{g}$, the median lies 0.0083 and the mean 0.013 above it, the median conductance lies 0.049 kW/K and the mean 0.10 kW/K above $\hat{G}$, and the exact standard deviation of $G$ is 0.591 kW/K, 5 % wider than the Laplace 0.561 kW/K. The skew matters most for the full-load question, whose event lies in the tail of $g$ toward large conductances, the side on which the exact posterior extends further: the Laplace tail 0.246 is low by 0.036 against the exact 0.281. Further into the tail the understatement grows: in Table 1.5 the Laplace tail on $g$ is smaller than the exact one by a factor of 2.5 to 4 in the two smallest rows, 0.0062 against 0.0155 and 0.0012 against 0.0049. Written on the first entry instead, as $T < T^* = 22.75$°C, the full-load event has the Laplace tail 0.253, because the normal follows the straight tangent rather than the curve; the exact value is 0.281 on either entry. The exact values rest on the base of equation (1.19), flat in $g$ along the curve; with a base flat in $T$ along the curve instead, the full-load value is 0.249, a reminder that an exact tail depends on the base measure as well as on the rows. The conductance band and the full-load probability are therefore slightly narrow and slightly low, and the report says so.

**A scalar check.** Eliminating $T$ through the physics row and differentiating the remaining one-entry objective in $g$ twice gives the same $\sqrt{\Sigma_{22}} = 0.0948$ and, through the slope of the

Table 1.8: The Laplace approximation of the posterior of the equipment hall against the exact posterior along the physics curve, in the limit $\sigma_p \to 0$.

| Quantity | Unit | Laplace, at $\hat{\mathbf{z}}$ | Exact, along the curve |
|---|---|---|---|
| Median of $T$ | °C | 23.070 | 23.028 |
| Mean of $T$ | °C | – | 23.030 |
| Standard deviation of $T$ | K | 0.4806 | 0.4813 |
| $\Pr(T < 24°\text{C} \mid \mathbf{y})$ | – | 0.973 | 0.977 |
| Median of $g$ | – | −0.0139 | −0.0056 |
| Mean of $g$ | – | – | −0.0013 |
| Standard deviation of $g$ | – | 0.0948 | 0.0967 |
| Median of $G$ | kW/K | 5.917 | 5.967 |
| Mean of $G$ | kW/K | – | 6.020 |
| Standard deviation of $G$ | kW/K | 0.561 | 0.591 |
| Median of $K_Q$ | K/kW | 0.1690 | 0.1676 |
| Mean of $K_Q$ | K/kW | – | 0.1677 |
| Standard deviation of $K_Q$ | K/kW | 0.0160 | 0.0160 |
| $\Pr(T_2 < 27.5°\text{C} \mid \mathbf{y})$ at 60 kW | – | 0.246 | 0.281 |
| $\Pr(T_2(17.1°\text{C}) < 27.5°\text{C} \mid \mathbf{y})$ | – | 0.605 | 0.639 |
| Largest set point at 0.95 | °C | 15.65 | 15.85 |

curve, the same $\sqrt{\Sigma_{11}} = 0.481\,\text{K}$. The vector form reaches them without the substitution, and it is the form that carries over unchanged to the cooling network of Section 1.1.3.

### 1.4.3 The residual check

Three rows act on two entries, and with the physics exact the reading and the handbook both inform the one direction along the curve: one row is redundant, and the rows at the MAP state measure how well the two agree. They are $r_1 = 0.005$, $r_2 = -0.260$ and $r_3 = -0.790$ (Table 1.4), and their squares sum to $2\mathcal{L}(\hat{\mathbf{z}}) = 0.692$. When the model and the standard deviations of the rows are right, that sum follows a chi-square distribution with as many degrees of freedom as there are rows less entries, here one; Chapter 26 returns to the result. A sum of 0.692 or more then has probability 0.41, while a sum above 3.84 would have probability below 0.05 and would say that the reading or the handbook value is wrong. The two agree, which is a first, coarse look at the question of a biased sensor.

### 1.4.4 The weight of the handbook prior

The handbook value is the one input of the fit that the hall did not supply, and removing its row shows how much the answers lean on it. Without the prior row the residual vector holds the rows $r_1$ and $r_2$ of equation (1.10): two rows, two entries, a square Jacobian, both rows zero at the MAP state (23.2°C, −0.0392), and $\Sigma_{11} = \sigma_T^2$ whatever $\sigma_p$, and $\sqrt{\Sigma_{22}} = 0.0962$ (the last exercise of the chapter derives both). On the conductance entry, the route of equation (1.33), the full-load probability is then

$$\Pr(T_2 < 27.5°\text{C} \mid y_T) \approx \Phi\left(\frac{-0.0392 - 0.0513}{0.0962}\right) = \Phi(-0.941) = 0.173,$$

against 0.246 with the handbook row, a rise of 0.073. Exactly, with the same base flat in $g$ along the curve and the prior term removed from equation (1.19), it is 0.212 against 0.281, a rise of 0.069. Panel (a) of Figure 1.7 draws the two Laplace marginals of the conductance entry.

The horizontal axis is $g$ and the vertical axis the probability density; the dashed curve is the marginal under the reading alone, $\mathcal{N}(-0.0392,\ 0.0962^2)$, and the solid curve the marginal with the handbook row, $\mathcal{N}(-0.0139,\ 0.0948^2)$. The dotted red line is the threshold $g^* = 0.0513$, and the shaded areas to its right are the two probabilities. The handbook value, which places the conductance above what the reading alone implies, moves the marginal toward the threshold with nearly the same width. Moving the set point moves the threshold of that panel, and panel (b) reads the result. Its horizontal axis is the set point $r$ and its vertical axis the probability of holding 27.5°C at 60 kW; the marks are the rows of Table 1.6, solid for the state route and dashed for the exact value. The horizontal dashed line is the level 0.95, and the dotted vertical line the coil's floor at 17.75°C, the limit stated in Section 1.6.5. The state route lies below the exact one at every set point, the gap narrowing to 0.015 at 15.5°C. The level 0.95 lies far to the left of the floor.

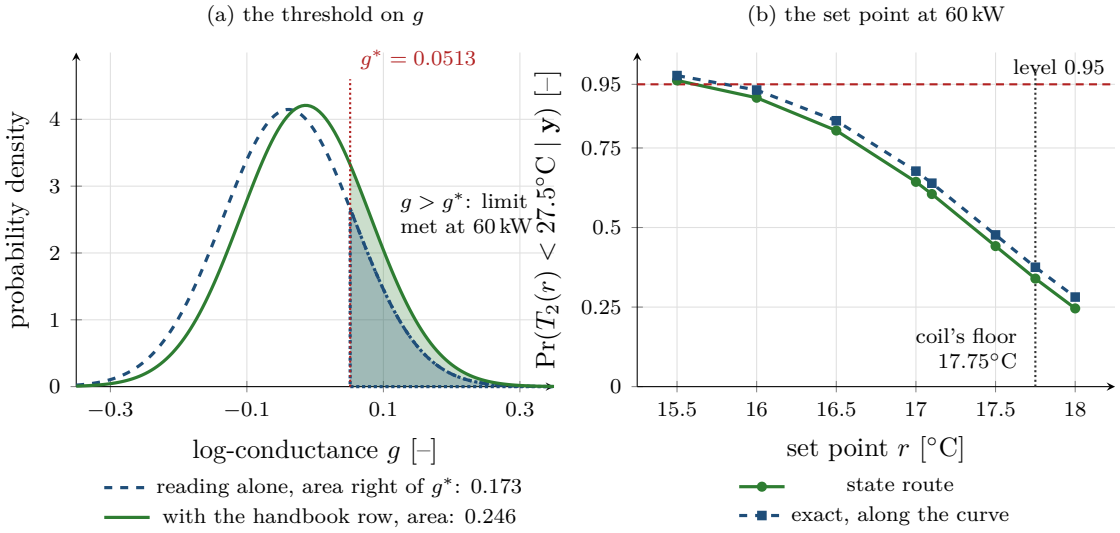



Figure 1.7: The handbook prior and the set point in the full-load answer. (a) Laplace marginals of the conductance entry, $\mathcal{N}(-0.0392,\ 0.0962^2)$ under the reading alone (dashed) and $\mathcal{N}(-0.0139,\ 0.0948^2)$ with the handbook row (solid), densities $\varphi((g-\hat{g})/s)/s$ with $\varphi$ the standard normal density. The shaded areas right of $g^* = 0.0513$ are the full-load probabilities, 0.173 and 0.246. (b) The probability of holding 27.5°C at 60 kW against the set point, on both routes (Table 1.6), with the level 0.95 and the coil's floor at 17.75°C, the stated limit of Part II; no set point left of the floor can be held at 60 kW, and the marks there read the question at a pinned supply air.

The same holds in every row of Table 1.5, on both routes: without the handbook row each probability is lower than with it. The full-load answer therefore depends on the handbook value, and the handbook value is stated with the report.

## 1.5 The report (step 6)

At a pinned supply air, the commissioning sheet's 17.1°C holds the hall with probability only 0.61 (0.64 exact), not "holds", and holding at 0.95 takes 15.65°C, if the coil could deliver it; it stops at 17.75°C, 0.34 (0.38 exact). The handbook conductance of 7.5 kW/K, carried forward, put the hall at 26°C at 60 kW, well below the limit of 27.5°C; held against the one reading through the energy balance, it gives way to a conductance of 5.92 kW/K, below the 6.316 kW/K the full load needs, and at the present set point the hall stays below 27.5°C at 60 kW with probability of about one in four, 0.246 by the Laplace route and 0.281 exactly: neither the

handbook's nothing to do nor the sheet's fixed correction. What the loop changes is that the threshold can be moved: the set point adds no spread, and every set point the coil can reach has its probability. A coil heat-removal check, the reading that seems to measure the very heat in question, would not change that answer (Section 1.6.2). The answers to the questions of Table 1.2 are given below in the units in which they were asked, and the decision on the owner's request follows from the last two.

### 1.5.1 The answers and their bands

Table 1.9 gives the answers. Each row is one question; the columns give the answer computed from $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$, which the chapter reports, the exact value beside it from the density along the curve, and the band (one standard deviation for the temperature and the process gain, Laplace and exact; for the full-load probability, its value without the handbook row, the check of Section 1.4.4; for the set point, the value at the coil's floor). The added-sensor question is answered in Section 1.6.

Table 1.9: The report on the equipment hall from the reading $y_T = 23.2°\mathrm{C}$ and the handbook conductance of 7.5 kW/K, with the physics exact.

| Question | Answer (state route) | Exact | Band |
|---|---|---|---|
| What temperature is the hall running at now? | $\hat{T} = 23.07°\mathrm{C}$ | median 23.03°C | ±0.48 K (0.481 K exact) |
| What is the process gain to the load? | $\hat{K}_Q = 0.1690\,\mathrm{K/kW}$ ($\hat{G} = 5.92\,\mathrm{kW/K}$) | median 0.1676 K/kW (mean 0.1677 K/kW) | ±0.0160 K/kW (0.0160 K/kW exact); on $G$, ±0.56 kW/K (0.59 kW/K exact) |
| Is the hall within its 24°C limit now? | $\Pr(T < 24°\mathrm{C} \mid \mathbf{y}) =$ 0.973, about 0.97 | 0.977; gap 0.004 | – |
| Can the hall carry 60 kW within 27.5°C at this set point? | $\Pr(T_2 < 27.5°\mathrm{C} \mid \mathbf{y}) =$ 0.246, about 0.25 | 0.281; gap 0.036 | without the handbook value 0.173 (0.212 exact) |
| How far must the set point move to hold it? | at a pinned supply air, 0.605 at 17.1°C; 0.95 at 15.65°C | 0.639; 15.85°C | neither is reachable: the coil's floor 17.75°C gives 0.340 (0.375 exact) |

The reading has placed the hall temperature to within about half a kelvin, slightly better than the sensor's own 0.5 K because the handbook value adds a little through the physics row. The conductance is known to within 9.5 % of its value, against a factor $\mathrm{e}^{0.3} = 1.35$ in the handbook before the reading; the exact posterior widens that band to 0.59 kW/K. The process gain is 0.1690 K/kW, not the sheet's 0.1733 K/kW. The hall is below 24°C with probability of about 0.97.

### 1.5.2 The full-load decision

The owner's request has the answer of the fourth row: at 60 kW the hall stays below 27.5°C with probability of about one in four, 0.246 by the Laplace route and 0.281 exactly: on the present cooling the record does not support doubling the load. The answer assumes that the loop still holds the supply air at 18°C at 60 kW; if it cannot, $T_c$ rises with the load and the hall is warmer than this answer says (Section 1.6.5). The reason is read from the equivalent threshold: the limit is met only if the conductance exceeds $G^* = 6.316\,\mathrm{kW/K}$, and the posterior places the conductance at $5.92 \pm 0.56\,\mathrm{kW/K}$; in the conductance entry the threshold $g^* = 0.0513$ lies 0.69 standard deviations above $\hat{g} = -0.0139$. The answer leans on the handbook value:

without it the probability would be 0.173 by the Laplace route and 0.212 exactly. A reading cannot raise the conductance; it can only show where it lies, and with the estimate below the threshold, further evidence is more likely to confirm the shortfall than to remove it. The set point alone does not settle it either: the probabilities of the fifth row hold only down to the coil's floor (Section 1.6.5), and at that floor, with the coil at its rating, the set point reaches 0.34 (0.38 exact). Running the hall at 60 kW under this limit calls for more cooling: more airflow (a larger conductance) or a coil that can deliver a colder supply air; not more readings.

### 1.5.3 The entry in the commissioning log

The engineer's entry reads the fit and not the sheet. The hall runs at $23.07 \pm 0.48$°C under 30 kW, on one sensor and the handbook value, with the discharge loop holding 18°C. The process gain to the load is $0.1690 \pm 0.0160$ K/kW. At 60 kW and a pinned supply air of 17.1°C the hall holds 27.5°C with probability 0.605 (0.639 exact), and 0.95 takes 15.65°C, so the sheet's correction is not entered as a fix. Neither is reachable on this coil: below 17.75°C the supply air stays at the floor, 0.34 (0.38 exact). The request to double the load is referred for more cooling. The next measurement is a second sensor in the hall, with an error independent of the first (Section 1.6.3).

## 1.6 The next measurement (step 7)

Whether the bands of Table 1.9 are narrow enough is the engineer's decision. If they are not, the added-sensor question of Table 1.2 asks which reading to add to narrow them most, and a reading not yet taken is a row not yet attached.

### 1.6.1 The preposterior covariance

The value a candidate reading will return is unknown, but its Jacobian row at $\hat{\mathbf{z}}$, its sensitivity to each entry divided by its standard deviation, is known before the reading is taken, and that is enough. The information matrix is a sum of one outer product per row (Section 1.2.1), so a candidate row with the Jacobian row $\mathbf{J}_a$ adds $\mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a$ to the inverse of the present covariance, and the covariance expected after the reading is

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1}, \tag{1.37}$$

with $\boldsymbol{\Sigma}$ of equation (1.21), curvature term included. The value the reading will return does not appear in it. Chapter 2 derives $\boldsymbol{\Sigma}_a$ as the preposterior covariance, equation (2.17), and Chapter 23 ranks candidate tests by it in excitation design. The band on the temperature after the reading is $\sqrt{\Sigma_{a,11}}$, and the band on the conductance, the quantity on which the process gain and the full-load answer rest, is $\sigma_{G,a} = \sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_a\mathbf{g}}$ with the gradient of equation (1.24). The two gradients (1.24) and (1.25) act on the same entry, so a candidate removes the same fraction from the band on the gain as from the band on the conductance.

### 1.6.2 The candidate readings

Four candidates are considered, each written as a row over $\mathbf{z} = (T, g)^{\mathsf{T}}$:

- a second hall sensor of the same accuracy, with an error independent of the first: the row $(T - y'_T)/\sigma_T$, with the Jacobian row $\mathbf{J}_a = (2,\ 0)$, the same as the present reading row;
- a reference sensor of 0.2 K in the hall: $\mathbf{J}_a = (5,\ 0)$;
- a reading of the supply-air temperature, or a load meter on the servers: $T_c$ and $Q$ are pinned, not entries of the state, so the row has no entry in either column, $\mathbf{J}_a = (0,\ 0)$;
- a coil heat-removal check (the airflow through the unit times the temperature drop across its coil), $\sigma_q = 1$ kW on the heat removed $G\,(T - T_c)$: $\mathbf{J}_a = \left(\hat{G},\ \hat{G}\,(\hat{T} - T_c)\right)/\sigma_q = (5.917,\ 30.00)$.

Table 1.10 evaluates equation (1.37) for each. Each row is one candidate; the columns give its Jacobian row, the standard deviation of $T$ after it, $\sqrt{\Sigma_{a,11}}$, and the width it removes from the present 0.481 K, then the conductance band after it, $\sigma_{G,a}$ with the gradient $\mathbf{g} = (0,\ 5.917)^{\mathsf{T}}$, and the width it removes from the present 0.561 kW/K. One row of the arithmetic, for the second sensor: $\mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a$ has the single entry 4, in its $T$–$T$ position, and

$$\boldsymbol{\Sigma}_a = \begin{bmatrix} 0.12005 & -0.023678 \\ -0.023678 & 0.0046738 \end{bmatrix}, \qquad \sqrt{0.12005} = 0.346\,\text{K}, \quad 5.917\,\sqrt{0.0046738} = 0.405\,\text{kW/K}.$$

Table 1.10: Candidate readings for the equipment hall ranked by the preposterior covariance of equation (1.37), evaluated at $\hat{\mathbf{z}}$. Present bands: 0.481 K on $T$ and 0.561 kW/K on $G$.

| Candidate | $\mathbf{J}_a$ [$\text{K}^{-1}$, –] | $\sqrt{\Sigma_{a,11}}$ [K] | removed [K] | $\sigma_{G,a}$ [kW/K] | removed [kW/K] |
|---|---|---|---|---|---|
| Second hall sensor, 0.5 K | (2, 0) | 0.346 | 0.134 | 0.405 | 0.156 |
| Reference hall sensor, 0.2 K | (5, 0) | 0.185 | 0.296 | 0.216 | 0.345 |
| Supply-air reading or load meter | (0, 0) | 0.481 | 0 | 0.561 | 0 |
| Coil heat-removal check, 1 kW | (5.917, 30.00) | 0.481 | $< 10^{-6}$ | 0.561 | $< 10^{-6}$ |

Only the two sensors in the hall remove anything, and the reason is visible in their Jacobian rows. The posterior is wide in one direction only, along the physics curve, $(0.9811,\ -0.1935)^{\mathsf{T}}$, where the present information is 4.17: the 4.27 of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ less the 0.099 of the curvature term. The second sensor's row projects on that direction as $2 \times 0.9811 = 1.962$ and adds $1.962^2 = 3.85$, nearly doubling the information, so that the variances along the curve fall by the factor $4.17/8.02 = 0.52$ and both standard deviations by the factor 0.72. The supply-air reading and the load meter read pinned quantities: in this model they are known already, their rows are zero, and they would inform only a graph in which $T_c$ or $Q$ were free. The heat-removal check is the subtler case. Its row (5.917, 30.00) is proportional to the physics row (100, 507.0), both in the ratio 5.070, so it points across the curve, where the physics already fixes the state to within $\sigma_p$; it measures again what the energy balance asserts, and it adds nothing along the curve. This row is the product alone; an airflow reading by itself would be a different row, on $g$, and does not fix $G$ because $G$ also lumps mixing. A reading is worth what it adds in the direction in which the posterior is wide.

### 1.6.3 The recommendation

The recommendation, and the answer to the added-sensor question, is a second sensor in the hall, a permanent instrument: it removes 0.134 K of the 0.481 K on the temperature and 0.156 kW/K of the 0.561 kW/K on the conductance, 28 % of each, and a 0.2 K reference sensor, if one is available, removes 62 %. A supply-air reading, a load meter and a coil heat-removal check are not worth requesting for these answers. A load step is a different kind of test, one the steady graph of this chapter cannot value, and Chapter 23 values it. Two conditions apply. The second sensor's error must be independent of the first: a second reading from the same sensor at the same condition would share any bias of the first, and a shared error is a discrepancy that repeats (Section 1.1.4), which a repeated reading does not reduce. And the gain is computed at $\hat{\mathbf{z}}$ under the Laplace approximation, which the checks found adequate for $T$ and slightly narrow for $G$. Once the reading is taken, its row joins the three of equation (1.10), and the method returns to

step 3 with four rows. For the full-load decision the narrower conductance band sharpens the answer but does not move the threshold: the decision of Section 1.5.2 turns on more cooling, more airflow or a coil that can deliver a colder supply air, not on more readings.

### 1.6.4 What the analysis showed

The handbook, carried forward, put the hall at 26°C at 60 kW, with nothing to do, and the commissioning sheet, taking the reading as the truth, at 28.4°C, with a set point of 17.1°C as the fix; the analysis, which holds the two against each other through the energy balance, finds that 17.1°C holds with probability 0.605 (0.639 exact) at a pinned supply air, which this coil cannot reach, its floor of 17.75°C giving 0.34 (0.38 exact), and that at the present set point the probability is about one in four that the hall stays below 27.5°C, 0.246 from the posterior of the state (0.281 exact). The loop's contribution is a threshold that can be moved at no cost in spread, down to the coil's floor, not a certainty: a flat trend at the set point left the hall's own uncertainty where it was. The analysis also overturns the obvious candidate: a coil heat-removal check, although it measures the very heat the question is about, removes nothing, because it repeats what the energy balance already asserts. One reading and a handbook value pin the hall temperature to about half a degree but the conductance of its air path only to about ten percent, and the two are not known separately: the posterior lives on the physics curve, a ridge along which a warmer hall is a hall with a weaker path, which is why the correlation of the two entries is $-0.9998$. The owner's question about the full load is a question about that conductance. The hall meets 27.5°C at 60 kW at this set point only if its path conducts more than 6.316 kW/K; the evidence places it below that, and the answer leans on the handbook value. What would narrow the answers is a second sensor in the hall, which removes 28 % of both bands. What would change the full-load answer is more cooling: more airflow (a larger conductance) or a coil that can deliver a colder supply air; not more readings.

### 1.6.5 Limits of the model

The posterior is exact only under the model, the steady-state graph of one hall with one path to a supply air that the discharge loop holds, and that model can be wrong before any inference is run. The supply air is pinned only while the controller holds it: a valve at its stop can no longer hold $T_c$, which then rises with the load and would have to become a free quantity of the graph, including at the 60 kW of the scenario node, which the full-load answer assumes the loop can still hold at 18°C.

The set-point answer meets this limit directly. The coil of Part II has a conductance of 16 kW/K fully open, on chilled water at 14°C; this is a stated plant value, read here as a limit on the answer. Chapter 5 draws this coil, and Chapter 15 takes up the band at the stop. With that coil at its rating, and $T_w$ as stated, the valve is fully open at 60 kW at $r = 17.75$°C, and no lower set point can be held: below it the supply air stays at 17.75°C. The floor does not depend on $g$, since at steady state the coil removes $Q$ whatever the air path conducts. A coil below its rating raises the floor; no spread is carried on it here. As a plant sensitivity only, a coil log-conductance $\ln(W/W_{\text{rated}})$ of $-0.1$ or $+0.1$ puts the floor at 18.14 or 17.39°C, with probabilities 0.198 and 0.486 by the state route, and below $-0.0645$ the coil cannot hold even 18°C at 60 kW. The set point alone therefore reaches 0.34 (0.38 exact) in Table 1.6, and the 17.1°C of the sheet and the 15.65°C of the 0.95 level both lie beyond the coil. The decision stays more cooling.

The record itself is illustrative: the hall, the 23.2°C reading and the handbook value are constructed for the example, not field data, and no truth is stated and no comparison with a true state is claimed.

**Modeling mistakes.** Several errors can be made before any inference algorithm is run, in the construction of the graph and its rows. Table 1.11 lists five. Each row names the mistake, gives its form in the equipment hall, and states the correction.

Table 1.11: Five modeling mistakes made before any inference is run.

| Mistake | In the equipment hall | Correction |
|---|---|---|
| Treating a reading as the true state | the hall sensor reports 23.2°C, and $T = 23.2$°C is assumed | the reading is evidence with error $\sigma_T$; $T$ stays a free entry of the state |
| Confusing parameters and observations | the conductance and the reading are treated alike | $G$ is a physical property, an entry of the state; $y_T$ is evidence about $T$, connected to $G$ through the physics row |
| Assuming the physical model is complete | the steady-state relation is used just after a fan trip | stored heat matters in a transient, $C_{\mathrm{th}}\,\mathrm{d}T/\mathrm{d}t = Q - G\,(T - T_c)$ |
| Confusing structural with parameter uncertainty | doubt whether the hall has one or two effective air paths is folded into the spread of $G$ | the value of $G$ is a parameter question within one model; the number of paths is a question between two models |
| Believing the graph alone determines the posterior | a correct graph is expected to give a narrow answer | useful inference also needs suitable evidence, assumptions and constraints; some entries may remain poorly identified |

The first two mistakes confuse the roles of Table 1.3. A sensor reporting 23.2°C does not establish that the true temperature is exactly 23.2°C: the observation has an uncertainty, and the physics row of Section 1.1.4 showed that 23.2°C and 23°C imply conductances of 5.77 and 6 kW/K, a difference of 4 %. The conductance is a physical property of the path; the reading is evidence about the temperature, which is connected to the conductance through the model. The fourth mistake mixes two kinds of uncertainty that require different treatments. Uncertainty about the numerical value of $G$ within the model of Figure 1.2 is represented by the posterior of $g$. Uncertainty about whether the hall has one or two effective air paths is uncertainty about the model $M$ itself, and it is represented by comparing the evidence of the two drawings. The fifth mistake is the converse of the first. A graph provides relations; the posterior also requires evidence, and some entries may remain poorly identified even when the model is perfectly specified. In the two-path drawing, for example, a reading of the hall temperature determines only the total conductance of the two paths and leaves their split to the prior: the Jacobian of the physics and reading rows has a null vector along the split, and only a prior row adds information in that direction.

**The steady-state assumption.** The third mistake concerns the physics itself. The model of equation (1.1) assumes steady conditions and a single effective conductance. During transient operation the hall stores and releases heat, and the energy balance gains a storage term,

$$C_{\mathrm{th}}\,\frac{\mathrm{d}T}{\mathrm{d}t} = Q - G\,(T - T_c),$$

where $C_{\mathrm{th}}$ is the thermal capacitance of the hall, its air, structure and servers, in kJ/K, and $t$ is time in seconds. For constant $Q$, $G$ and $T_c$, write the departure from the steady temperature as $u = T - T_c - Q/G$. Then $\mathrm{d}u/\mathrm{d}t = \mathrm{d}T/\mathrm{d}t$ and the balance becomes

$$C_{\mathrm{th}}\,\frac{\mathrm{d}u}{\mathrm{d}t} = -G\,u,$$

whose solution is $u(t) = u(0)\,\mathrm{e}^{-t/\tau}$ with the time constant $\tau = C_{\mathrm{th}}/G$. Therefore

$$T(t) = T_c + \frac{Q}{G} + \Big(T(0) - T_c - \frac{Q}{G}\Big)\,\mathrm{e}^{-t/\tau}, \qquad \tau = \frac{C_{\mathrm{th}}}{G}. \tag{1.38}$$

The steady-state relation of equation (1.2) is the limit $t \to \infty$. Suppose (illustratively) that at $t = 0$ the cooling unit loses half its airflow (one of its two fans trips), so that the conductance falls from 6 to 3 kW/K while the hall is at 23°C under 30 kW. The new steady temperature is $18 + 30/3 = 28$°C, and equation (1.38) gives $T(t) = 28 - 5\,\mathrm{e}^{-t/\tau}$: after one time constant the hall is at $28 - 5\,\mathrm{e}^{-1} = 26.16$°C. A steady-state model would place it at 28°C at once. It is therefore inadequate for predicting temperatures immediately after a fan trip, where the time constant decides whether the engineer has minutes or seconds; the graph for the question of the lost conductance needs the storage term and $C_{\mathrm{th}}$ as a further entry or a pinned input.

The chapter has built the representation on which every question will be asked. The next chapter, Chapter 2, turns to the questions themselves: how a control engineer's question stated in the language of the loop is recognized as estimation, prediction, comparison, diagnosis or design, and written as a precise probability or expectation over the state of such a graph. From Part II the hall grows: the discharge loop is drawn, with its coil on chilled water, its valve and its controller.

## 1.7 What to retain

**Definition 1.1** (Plant graph)**.** A plant graph represents a physical system as free and pinned nodes, paths and sources, with the physical relations that must hold among them written as residual rows; its free quantities, physical states and parameters alike, form the state vector $\mathbf{z}$; readings, records and prior knowledge are attached as further rows; and every control engineer's question is then a probability or an expectation over $\mathbf{z}$, conditional on the evidence $\mathbf{y}$, computed from the posterior $p(\mathbf{z} \mid \mathbf{y})$ that the rows define.

The forward calculation, the posterior and the question are three different things computed over the same graph. The forward calculation uses no evidence and gives a number; the posterior combines the physics, the readings and the prior; the question picks what is read from the posterior. The posterior is computed in state-vector form: the rows, each divided by its standard deviation, form one residual vector over $\mathbf{z}$; the normal equations $(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}$ give the Gauss–Newton steps to the MAP state $\hat{\mathbf{z}}$; the information matrix $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ is a sum of one outer product per row; the covariance $\boldsymbol{\Sigma}$ is the inverse of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ plus the rows' own curvature, here the curvature of the physics curve; and an answer is an entry of $\hat{\mathbf{z}}$ with its $\sqrt{\Sigma_{ii}}$, or a derived quantity with $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}}$. An exact physical law is the limit $\sigma_p \to 0$ of its row, and its signature in $\boldsymbol{\Sigma}$ is a correlation near $-1$: a ridge along the zero set of the row, in the direction of the null vector of the row's Jacobian. A question about a quantity that is a monotone function of one entry of the state is transformed into a threshold on that entry by the physics alone, as the full-load question became $G > 6.316$ kW/K, and its probability is a normal tail of that entry's marginal. The answer is computed from $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$, with the exact value along the curve set beside it and the gap stated, checked against the full calculation (the limit $\sigma_p \to 0$, the exact posterior along the curve, the agreement of the redundant rows, the weight of the prior), reported with its band in the units of the question, and followed by the next reading, ranked by the preposterior covariance $\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$ of a row not yet taken; a reading of a pinned quantity, or a row parallel to a physics row, removes nothing. A process gain is a derived quantity like any other. A set point is a pinned input of a scenario node: its gain on the temperature is one, it adds no spread, and it moves only the threshold, down to the point at which the coil can no longer hold it. Table 1.12 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 1.12: The terms of the plant graph and where each is defined. Units for the equipment hall.

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| Plant graph | nodes, paths, sources, rows | – | the system before any question | 1.1.4 |
| Free, pinned | $T$, $g$ free; $T_c$, $Q$ pinned | – | inferred against known | 1.1.4 |
| Energy balance | $Q = G\,(T - T_c)$ | kW | conservation at the hall node | 1.1.4 |
| Forward calculation | $T = T_c + Q/G$ | °C | temperature from known inputs | 1.1.4 |
| State vector | $\mathbf{z} = (T, g)^{\mathsf{T}}$ | °C, – | all free quantities at once | 1.1.4 |
| Log-conductance | $g = \ln(G/G_{\mathrm{ref}})$ | – | keeps $G$ positive | 1.1.4 |
| Physics row | $(T - T_c - (Q/G_{\mathrm{ref}})\mathrm{e}^{-g})/\sigma_p$ | – | a relation that must hold, tolerance $\sigma_p$ | 1.1.4 |
| Model discrepancy | $r_{\mathrm{phys}}(\mathbf{z}) = \delta$ | kW | model error, not sensor error | 1.1.4 |
| Reading row | $(T - y_T)/\sigma_T$ | – | evidence about the state | 1.1.7 |
| Prior row | $(g - g_0)/\sigma_g$ | – | knowledge before the readings | 1.1.7 |
| Jacobian pattern | which row acts on which entry | – | a property of the graph | 1.1.7 |
| Scenario node | $T_2 = T_c + (Q_2/G_{\mathrm{ref}})\mathrm{e}^{-g}$ | °C | an unobserved condition sharing $g$ | 1.1.7 |
| Factor graph | entries and rows | – | the graph of the inference | 1.1.3 |
| Delta-method band | $\sigma_G^2 = \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$, $\mathbf{g} = (0, \hat{G})^{\mathsf{T}}$ | kW/K | spread of a derived quantity | 1.3.1 |
| Process gain | $K_Q = \partial T/\partial Q = \mathrm{e}^{-g}/G_{\mathrm{ref}}$ | K/kW | hall temperature per kilowatt of load | 1.3.1 |
| Set point at a scenario node | $T_2(r) = r + (Q_2/G_{\mathrm{ref}})\mathrm{e}^{-g}$, $\partial T_2/\partial r = 1$ | °C | moves the threshold, adds no spread | 1.3.3 |
| Ridge | $\mathbf{J}_1\mathbf{n} = 0$, $\boldsymbol{\Sigma} \to \mathbf{u}\mathbf{u}^{\mathsf{T}}/\lambda_{\parallel}$, $\lambda_{\parallel} = (n_1^2/\sigma_T^2 + 1/\sigma_g^2 + n_1 r_2/\sigma_T)/(n_1^2 + 1)$ | – | posterior along the physics curve | 1.2.1 |
| Exact posterior on the curve | $p(g \mid \mathbf{y})$, equation (1.19) | – | the limit $\sigma_p \to 0$ by quadrature | 1.2.1 |
| Threshold probability | $\Pr(T_2 < T_{\max} \mid \mathbf{y})$ | – | prediction under a threshold | 1.3.2 |
| Equivalent threshold | $G^* = Q_2/(T_{\max} - T_c)$ | kW/K | the same event on an entry | 1.3.2 |
| Posterior | $p(\mathbf{z} \mid \mathbf{y})$ | – | plausible states after the evidence | 1.2.1 |
| Residual vector | $\mathbf{r}(\mathbf{z})$, $\mathcal{L} = \frac{1}{2}\mathbf{r}^{\mathsf{T}}\mathbf{r}$ | – | all rows in standard deviations | 1.1.7 |
| Jacobian | $\mathbf{J} = \partial\mathbf{r}/\partial\mathbf{z}$ | varies | sensitivity of each row to each entry | 1.1.7 |
| Normal equations | $(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}$ | varies | the Gauss–Newton step to $\hat{\mathbf{z}}$ | 1.2.1 |
| MAP state | $\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} \mathcal{L}(\mathbf{z})$ | °C, – | the most probable state | 1.2.1 |
| Information matrix | $\mathbf{J}^{\mathsf{T}}\mathbf{J} = \sum_k \mathbf{J}_k^{\mathsf{T}}\mathbf{J}_k$ | varies | precisions of the rows, added | 1.2.1 |
| Laplace covariance | $\boldsymbol{\Sigma} = \left(\mathbf{J}^{\mathsf{T}}\mathbf{J} + \sum_k r_k\,\partial^2 r_k/\partial\mathbf{z}\,\partial\mathbf{z}^{\mathsf{T}}\right)^{-1}$ | varies | spreads and correlation of the entries | 1.2.1 |
| Preposterior covariance | $\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$ | varies | spreads after a row not yet taken | 1.6.1 |
| Time constant | $\tau = C_{\mathrm{th}}/G$ | s | transient omitted by steady state | 1.6.5 |

## 1.8 Exercises

**Exercise 1.1.** (a) Which statement best describes the role of the hall sensor: it defines the physical law; it reports the exact true state; or it supplies evidence about an uncertain state? (The third: the sensor provides noisy evidence, while the physical model relates the state to the inputs and the conductance.) (b) A hall loses heat to the same supply air through two cooling units' air paths, with conductances $G_1$ and $G_2$. Name the free and pinned quantities, write the physics row, and say what a temperature reading alone can determine. (Free: $T$, $g_1$, $g_2$; pinned: $T_c$, $Q$; the physics row, normalized as in equation (1.5), is $\big(T - T_c - Q/(G_{\mathrm{ref}}(\mathrm{e}^{g_1} + \mathrm{e}^{g_2}))\big)/\sigma_p$; on its zero set its Jacobian row over $\mathbf{z} = (T, g_1, g_2)^{\mathsf{T}}$ is $\big(1,\ (T-T_c)G_1/(G_1+G_2),\ (T-T_c)G_2/(G_1+G_2)\big)/\sigma_p$ and the reading row's is $(1/\sigma_T,\ 0,\ 0)$, so the two-by-three Jacobian of the two rows has the null vector $(0,\ G_2,\ -G_1)^{\mathsf{T}}$, the step $\Delta G_1 = -\Delta G_2$ that leaves the total unchanged; $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ is singular along it, the reading determines only $G_1 + G_2$, and the split rests on the priors.) (c) Write "will the hall stay below 27°C if the load rises to 45 kW?" as a probability, and say what the graph needs. ($\Pr(T_3 < 27^{\circ}\mathrm{C} \mid \mathbf{y})$, with a scenario node $T_3 = T_c + (45/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$ at 45 kW sharing $g$, unobserved.)

**Exercise 1.2.** (a) A room is cooled by supply air at 20°C; it has a heat load of 24 kW and a conductance of 4 kW/K. Compute its steady-state temperature. ($T = 20 + 24/4 = 26^{\circ}$C.) (b) A room is cooled by supply air at 19°C; it has a load of 40 kW and a strict limit of 27°C. What conductance keeps it below the limit? ($19 + 40/G < 27$ gives $40/G < 8$, so $G > 5$ kW/K.) (c) For the hall of the chapter, with $G_{\mathrm{ref}} = 6$ kW/K, evaluate $r_{\mathrm{phys}}$ and $r_{\mathrm{obs}}$ at $\mathbf{z} = (23.2^{\circ}\mathrm{C},\ 0)$ and at $\mathbf{z} = (23^{\circ}\mathrm{C},\ 0)$. (1.2 kW and 0; 0 and $-0.4$: the first state matches the reading and violates the physics, the second satisfies the physics and lies 0.4 standard deviations from the reading.)

**Exercise 1.3.** Suppose both the supply-air temperature $T_c$ and the conductance $G$ are uncertain. Starting from $T = T_c + Q/G$, write the probability that the temperature exceeds a limit $T_{\max}$, and state what joint distribution is needed to evaluate it. ($\Pr(T > T_{\max} \mid \mathbf{y}) = \iint \mathbf{1}\{T_c + Q/G > T_{\max}\}\, p(T_c, G \mid \mathbf{y})\, \mathrm{d}T_c\, \mathrm{d}G$, which needs the joint posterior of $T_c$ and $G$; if the two are independent given $\mathbf{y}$, it reduces to $\int p(T_c \mid \mathbf{y})\, \Pr(G < Q/(T_{\max} - T_c) \mid \mathbf{y})\, \mathrm{d}T_c$ over $T_c < T_{\max}$, plus $\Pr(T_c \geq T_{\max} \mid \mathbf{y})$.)

**Exercise 1.4.** For the two-condition hall of Figure 1.3, with a proposed load of $Q_2 = 50$ kW and a limit of 28°C: (a) derive the threshold conductance and the equivalent threshold on the present temperature. ($T_2 = 18 + 50/G < 28$ gives $G > 5$ kW/K; by equation (1.32), $T < 18 + (30/50) \times 10 = 24^{\circ}$C.) (b) Show that the present temperature determines the full-load temperature without knowledge of $G$. (On the physics curve both temperatures are functions of $g$, $T = T_c + (Q_1/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$ and $T_2 = T_c + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$, so a step along the null vector of the physics row moves $T_2$ by $Q_2/Q_1 = 5/3$ times the move of $T$; in closed form, dividing the two balances gives $T_2 - T_c = (Q_2/Q_1)(T - T_c)$, equation (1.29).) (c) If the reading $y_T = 23.2^{\circ}$C with $\sigma_T = 0.5$ K were the only information about $T$, what would the probability of meeting the limit be? ($\Phi((24 - 23.2)/0.5) = \Phi(1.6) = 0.945$.)

**Exercise 1.5.** The hall is at 23°C under 30 kW with $T_c = 18^{\circ}$C when its conductance falls from 6 to 3 kW/K; its thermal capacitance is $C_{\mathrm{th}} = 3600$ kJ/K (illustrative). (a) Compute the time constant and the new steady temperature. ($\tau = 3600/3 = 1200$ s $= 20$ min; 28°C.) (b) Compute the temperature after 10 and 60 minutes, and the time at which it reaches 27.5°C. ($28 - 5\,\mathrm{e}^{-0.5} = 24.97^{\circ}$C; $28 - 5\,\mathrm{e}^{-3} = 27.75^{\circ}$C; $5\,\mathrm{e}^{-t/\tau} = 0.5$ gives $t = \tau \ln 10 = 46.1$ min.) (c) State what a steady-state model would predict and when it is adequate. (It predicts 28°C immediately; it is adequate for conditions held for several time constants, not for the first minutes after the fan trip.)

**Exercise 1.6.** Remove the prior row from the hall of Section 1.2.1, so that the residual vector holds the rows $r_1$ and $r_2$ of equation (1.10) over $\mathbf{z} = (T, g)^{\mathsf{T}}$, with $\sigma_p = 0.01$ K and

$\sigma_T = 0.5\,\mathrm{K}$. (a) Find the MAP state and write the two-by-two Jacobian there. (Both rows vanish at $T = y_T = 23.2°\mathrm{C}$ and $g = \ln(5/5.2) = -0.0392$; there $\mathbf{J} = \left[\begin{smallmatrix} 100 & 520 \\ 2 & 0 \end{smallmatrix}\right]$, since $5\,\mathrm{e}^{0.0392}/0.01 = 520$.) (b) Show that $\mathbf{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ has $\Sigma_{11} = \sigma_T^2$ for every $\sigma_p$, and compute $\sqrt{\Sigma_{22}}$. (The Jacobian is square and invertible, so $\mathbf{\Sigma} = \mathbf{J}^{-1}\mathbf{J}^{-\mathsf{T}}$. Writing $\mathbf{J} = \left[\begin{smallmatrix} a & b \\ c & 0 \end{smallmatrix}\right]$, $\mathbf{J}^{-1} = \left[\begin{smallmatrix} 0 & 1/c \\ 1/b & -a/(bc) \end{smallmatrix}\right]$; its first row gives $\Sigma_{11} = 1/c^2 = \sigma_T^2 = 0.25\,\mathrm{K}^2$, and its second $\Sigma_{22} = (\sigma_p^2 + \sigma_T^2)/(5\,\mathrm{e}^{-\hat{g}})^2$, so $\sqrt{\Sigma_{22}} = \sqrt{0.2501}/5.2 = 0.0962$.) (c) Explain why no curvature term is needed here, and how the result gives the reading-alone probability 0.173 of Section 1.4.4. (Both rows are zero at $\hat{\mathbf{z}}$, so $\sum_k r_k\,\partial^2 r_k/\partial\mathbf{z}\,\partial\mathbf{z}^\mathsf{T}$ vanishes; the entry $g$ then has the Laplace marginal $\mathcal{N}(-0.0392,\ 0.0962^2)$, and its tail above $g^* = 0.0513$ is $\Phi((-0.0392 - 0.0513)/0.0962) = \Phi(-0.941) = 0.173$.) (d) The reading alone gives the process gain $K_Q = 0.1733\,\mathrm{K/kW}$. Explain why the fit of the chapter gives a smaller value. (The handbook row, with $g_0 = 0.2231$ above the reading's $-0.0392$, pulls $g$ up to $\hat{g} = -0.0139$, a larger conductance and so a smaller gain: $K_Q = \mathrm{e}^{0.0139}/6 = 0.1690\,\mathrm{K/kW}$.)

# Chapter 2

# The Language of Control Questions

## 2.1 Control questions and their mathematical form

Chapter 1 drew the equipment hall as a plant graph: one hall node whose air temperature balances a heat load against one heat-removal path to the supply air, which the discharge loop holds at its set point, a pinned 18°C, with the conductance of that path unknown and a temperature sensor attached to the hall. The graph was built before any question was asked of it. At 19:00, the start of the night run, the commissioning sheet shows the hall sensor at 23.2°C, 0.8 K inside the 24°C limit, the trend of the discharge loop lies flat on its set point, and a work order asks for the load to go from 30 to 60 kW tonight; on the sheet alone the hall is in band and the loop is healthy. The engineer has one commissioning reading, the one process gain taken from it and a loop trend, and four questions on the desk: what the gain is, what the probability is that 60 kW holds, whether a rise in the reading is the path or the sensor, and whether a step test is worth a slot in the commissioning schedule. The log of the loop's trips, on the next page, adds a fifth, whether the loop will run through the night without a trip, and a process gain that may have drifted since the loop was tuned adds a sixth, whether to retune it before the night peak. They are not the same kind of question even though they concern the same plant. The sheet gives each of them a number, and an inference algorithm can return a mathematically correct result to a question nobody asked; the step that prevents both is the translation of the engineer's question into the mathematical quantity that answers it. The question of this chapter is

*how a control engineer's question about a running plant is translated into a mathematical expression that probabilistic inference can evaluate, and how that expression determines what must be computed.*

The answer is short. Once a probability statement names its event, its evidence and its model, every question about the plant is a probability or an expectation over the state of its plant graph, with the unknown on the left of the conditioning bar and the evidence on the right. Control questions about a plant then fall into five categories, estimation, prediction under a threshold, hypothesis comparison, diagnosis and design, each with its general forms, integrals or ratios written without reference to any particular plant. Five computations supply every quantity those forms contain: the posterior of the state; the trend, which reads that posterior along time; the evidence of a model; the preposterior of a measurement not yet taken; and the expected consequence of an action. Two of them, the trend and the expected consequence, have general forms of their own, which the categories call on. What differs from one plant to the next is the specialization: which entry or function of the state the form is applied to, which event or family of models it names, and which objects of the fit carry the answer.

**Jargon: event, conditional probability, indicator function, valve stiction, loop trip, trip log, trip rate, step test, posterior predictive, trend, crossing time, evidence, Bayes factor, expected consequence, risk, value of information**

An *event* is a statement about the plant that is either true or false, such as "the hall temperature exceeds 28°C"; a probability assigns it a number between zero and one. *Valve stiction* is the static friction that holds a valve stem still until the controller's demand has built up enough to make it jump. A *loop trip* is the event that sends a loop from automatic to manual, in the trip log of Section 2.4.3, because the discharge valve has stuck; the *trip log* is the record of the times at which the loop tripped, and the *trip rate* is the expected number of trips per hour of logged operation. A *step test* is a deliberate step of a loop's set point or actuator, made to read the plant's response. A *conditional probability* $\Pr(A \mid B)$ is the probability of event $A$ when the information $B$ is treated as known. The *indicator function* $\mathbf{1}_A(\mathbf{z})$ of an event equals one for the states at which the event is true and zero elsewhere, so that the probability of the event is the expectation of its indicator. A *posterior predictive* probability is the probability of a future event that stays random even when the state is known, averaged over the posterior of the state. A *trend* is the change per unit time of a quantity the state carries, and the *crossing time* is the time at which a rising quantity reaches a stated level from below. The *evidence*, or marginal likelihood, of a model is the probability of the observed evidence averaged over the prior of that model's unknowns; the *Bayes factor* is the ratio of the evidences of two models. The *expected consequence* of an action, also called its expected loss, is the probability-weighted average of the consequences that may follow it; summed over the events that may follow, it is each event's probability times its consequence, the *risk* of the action. The *value of information*, here the expected value of sample information (EVSI), is the reduction in expected loss that a measurement is expected to bring because the decision can be taken after its result is known.

**Key idea**

Every control question about a plant is answered by a specific mathematical object computed from the same plant graph and its rows: a marginal, a tail probability, a predictive average, the posterior read along time, a ratio of evidences, a preposterior or an expected consequence minimized over actions. Five computations produce these objects, the posterior, the trend, the evidence, the preposterior and the expected consequence. The question, not the loop trend or the inference tool, decides which object, and so which computation, is required.

The questions are put to one plant, the equipment hall of Chapter 1, and to the record of its loop's trips, so that the differences between them lie in the questions alone.

### 2.1.1 The plant graph of the equipment hall

Figure 2.1 draws the plant graph of Chapter 1, with the scenario node that the hall example of Section 2.4.1 adds to it. The upper green box is the free node of the hall at the present load, whose temperature $T_1$ is an entry of the state; the lower green box is the same hall at the requested full load, a scenario node whose temperature $T_2$ is also an entry and which has no sensor; the gray box on the right is the pinned node, the supply air at $T_c = 18°\mathrm{C}$, the set point at which the discharge loop holds it; the two solid green arrows are the heat-removal path at the two loads, which carry one and the same effective thermal conductance $G$, held in the state as the log-conductance $g = \ln(G/G_{\mathrm{ref}})$; the white boxes are the sources, the heat loads $Q_1$ and $Q_2$; and the dashed box on the left, joined to the hall by a dashed arrow, is the observation, the reading $y_T$ of the present-load temperature. Each hall node carries one physical row, its

steady-state energy balance, written for a node of temperature $T$ and load $Q$ as

$$Q = G\,(T - T_c), \tag{2.1}$$

where $Q$ is the heat load in kW, $G$ the conductance in kW/K, $T$ the hall air temperature and $T_c$ the supply-air temperature, both in °C. The supply air is pinned because the discharge loop holds it at its set point, and this part of the book treats that controller as ideal; the pin fails when the valve reaches its stop, at a load its coil cannot carry, or when the loop leaves automatic. The hall examples assume neither; the trip log counts the second without modelling its effect on $T_c$. Solved for the hall temperature, it gives the relation used throughout the chapter,

$$T = T_c + \frac{Q}{G}. \tag{2.2}$$

The relation is deterministic: once $G$ is known, $T$ is known. What makes the hall a probabilistic problem is that $G$ is not known, and that a reading $y_T$ differs from $T$ by a measurement error.

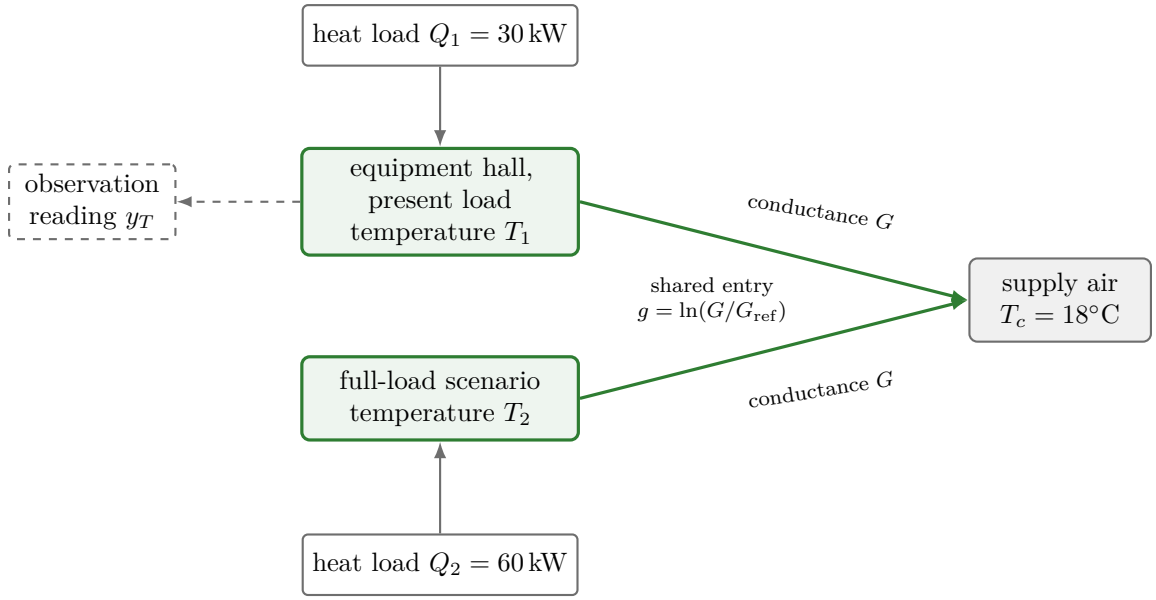



Figure 2.1: The plant graph of the equipment hall with the full-load scenario node. Free: the present-load temperature $T_1$ and the full-load temperature $T_2$ (green nodes) and the log-conductance $g$ shared by the two heat-removal paths (solid green edges). Pinned: the supply-air temperature $T_c = 18°\mathrm{C}$, the discharge loop's set point (gray node). Sources: the heat loads $Q_1 = 30\,\mathrm{kW}$ and $Q_2 = 60\,\mathrm{kW}$. Observed: the reading $y_T$ of the present-load temperature (dashed box and arrow); the scenario node has no sensor.

### 2.1.2 Four questions about one plant

The engineer's four questions are the rows of Table 2.1. Each row is a question in the engineer's words; the second column names the kind of calculation it requires and the third column the quantity that answers it, in the notation that Section 2.2 defines.

Question A estimates an unknown physical quantity. Question B asks for the probability of an event at a load not yet applied. Question C distinguishes between two explanations of the same readings, one in which the air path has changed and the sensor reads the hall temperature with zero-mean error, and one in which the path is unchanged and the sensor carries a bias. Question D predicts how a test that does not yet exist would change a commissioning decision. All four concern the same graph, and all four are computed from the same physical model; none of them is answered by the same calculation as another.

Table 2.1: Four questions about the equipment hall and the calculation each requires.

| Control engineer's question | Calculation | Answering quantity |
|---|---|---|
| A. What is the hall's process gain to the load, that is, the conductance of its air path? | estimation of an unknown physical quantity | $p(G \mid \mathbf{y})$ |
| B. With what probability does the hall hold below 27.5°C at 60 kW? | probability of an event | $\Pr(T_2 < 27.5^\circ\mathrm{C} \mid \mathbf{y})$ |
| C. Is a rise in the hall reading the path or the sensor? | comparison of competing explanations | $\Pr(M_{\text{path}} \mid \mathbf{y})$ against $\Pr(M_{\text{bias}} \mid \mathbf{y})$ |
| D. Is a step test worth a slot in the commissioning schedule? | prediction of what a future reading would change | the preposterior, and the expected loss after the test |

**Principle 2.1.** *The same physical graph does not imply the same probabilistic calculation. The formulation of the question determines which calculation is needed.*

## 2.2 The grammar of probability statements

A question can be classified only when the statement that expresses it is complete: an event or an unknown quantity, the evidence it is conditioned on, and the model under which its probability is computed.

### 2.2.1 Events

An event is a statement that can be true or false. Three events about the plant are

$$T > 28^\circ\mathrm{C}, \qquad G < 5\,\mathrm{kW/K}, \qquad N_a \geq 2,$$

where $N_a$ is the number of trips of the discharge loop during a night run. A probability assigns to an event a number between zero and one. The statement

$$\Pr(T > 28^\circ\mathrm{C}) = 0.10$$

means that the probability assigned to the temperature-exceedance event is 10 %. The statement is incomplete in one respect: it does not say how the probability was calculated or what information it incorporates. A probability computed before any reading of the hall was taken and one computed after thirty readings are different numbers for the same event, and the notation must say which is meant. That is the role of conditioning.

### 2.2.2 Conditional probability and the evidence

The probability of event $A$ given the information $B$ is written

$$\Pr(A \mid B), \tag{2.3}$$

and read "the probability of $A$, given $B$". In control the information on the right of the bar is the evidence: the loop trends and the telemetry of the present record (temperatures, loads, meter readings), the trip log, the results of step tests, the set points and run states of the equipment, nameplate and handbook values, and what the commissioning sheet established.

The book collects all of it in the bold symbol $\mathbf{y}$, and keeps a plain $y$ with a subscript, such as $y_T$, for a single reading. The statement

$$\Pr(T > 28^\circ\mathrm{C} \mid \mathbf{y})$$

asks for the probability that the hall temperature exceeds 28°C, conditional on the available evidence. It differs from $\Pr(T > 28^\circ\mathrm{C})$, which does not condition on that evidence explicitly. The unknown, the event about the plant, stands on the left of the bar; everything treated as known stands on the right. Keeping that order fixed is what prevents questions from becoming entangled when several are asked of one plant.

### 2.2.3 The role of the model

A complete probability statement also names the model under which it is computed:

$$\Pr(A \mid \mathbf{y}, M), \tag{2.4}$$

where $A$ is the event of interest, $\mathbf{y}$ the evidence, and $M$ the assumed mathematical model, which comprises the physical relations of the plant graph and the probability assumptions about errors and unknowns. The thermal model of equation (2.2) assumes steady state; a model that adds the thermal storage of the hall would assign a different probability to the same temperature event under the same evidence, because it predicts a different temperature after a change of load. The book sometimes suppresses $M$ to keep the notation readable. The dependence on the model remains in every probability, and a probability is a statement about the plant only to the extent that the model is.

### 2.2.4 Unknown quantities that are not events

An unknown conductance $G$ is a numerical quantity, not an event; it has a value, not a truth value. What can be asked of it is its probability distribution, described for a continuous quantity by a probability density,

$$p(G \mid \mathbf{y}, M),$$

where the lowercase $p$ denotes a density. Alternatively, a question can ask for the probability that the conductance falls within a range, for example

$$\Pr(5 < G < 7 \mid \mathbf{y}, M),$$

and that is an event probability, because "$5 < G < 7$" is a statement that is true or false. The two are connected by integration of the density over the range that defines the event:

$$\Pr(5 < G < 7 \mid \mathbf{y}, M) = \int_5^7 p(G \mid \mathbf{y}, M)\,\mathrm{d}G, \tag{2.5}$$

with the limits in kW/K. A distribution describes uncertainty across all possible values of a quantity; an event probability describes how much of that probability falls within a specified set of values. A density is not a probability: it has units, $\mathrm{kW}^{-1}\,\mathrm{K}$ here, and it can exceed one. Only its integral over a set is a probability.

### 2.2.5 A reusable notation

The questions of this chapter, and of the book, are written with the symbols of Table 2.2. The first column is the symbol, the second its meaning. The state $\mathbf{z}$ holds every free quantity of the plant graph at once, the physical quantities such as the log-conductance as well as the temperatures; there is no separate parameter vector. The action $a$ stands for a control decision

Table 2.2: The reusable notation of probabilistic questions.

| Symbol | Meaning |
|---|---|
| $\mathbf{z}$ | state vector of the plant graph: every unknown quantity, physical parameters included |
| $\mathbf{y}$ | observed evidence: readings, counts, records, priors |
| $M$, $M_j$ | assumed model or system structure; one of several competing models |
| $p(\mathbf{y} \mid M)$ | evidence of model $M$: the probability of the evidence with the model's state integrated out |
| $B_{AB}$ | Bayes factor of model $M_A$ against $M_B$, $p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$; $B_{10}$ is $M_1$ against $M_0$ |
| $A$ | an event |
| $f(\mathbf{z})$, $f^*$ | a quantity derived from the state; a requirement or limit on it |
| $\kappa$ | assurance level: the posterior probability a requirement must reach |
| $y_{\text{new}}$ | a future observation that stays random when the state is known |
| $a$ | a control action or decision, or a candidate reading |
| $\tilde{y}_a$ | the outcome of candidate reading $a$, not yet observed |
| $u$, $U(a)$ | a quantity of the posterior after $\tilde{y}_a$; its expectation over the outcomes, which ranks the candidates |
| $\mathbf{\Sigma}_a$ | covariance of the state after candidate measurement $a$ (the preposterior) |
| $L(a, \mathbf{z})$ | loss or cost of action $a$ when the state is $\mathbf{z}$: its consequence |
| $\mathcal{R}(a)$ | expected loss of action $a$ over the posterior, $\mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$ |
| $c_a$, $C_e$ | the fixed cost of action $a$; the consequence of event $e$ |
| $\mathbf{g}_a$, $\mathbf{H}_a$ | gradient and curvature of $L(a, \mathbf{z})$ in the state at $\hat{\mathbf{z}}$ |
| $f_t(\mathbf{z})$, $\mathbf{g}(t)$ | a quantity at time $t$ carried by a state indexed by time; its gradient |
| $\alpha$, $\gamma$ | the level and the trend entries of such a state: the quantity at the record's center $\bar{t}$ and its change per unit time (in a chance constraint, $\alpha$ is the stated risk) |
| $t^*$ | crossing time: the time at which a rising $f_t$ reaches a level $f^*$ |

such as retuning a loop, changing a set point, taking a loop to manual or running a step test; the symbol $\tilde{y}_a$ stands for the outcome of a candidate reading $a$ that has not yet been taken, and $y_{\text{new}}$ for a future outcome of the plant itself, such as the time of the next trip.

With these symbols, questions as different as the conductance of an air path, a clean night run and the choice of a step test are written in one mathematical language, and the differences between them become visible as differences in the expression, not in the words.

## 2.3 The probability over the state, the general forms and the five computations

Every question of Section 2.1 is an integral over the posterior of the state, a ratio of such integrals, or the minimum of one over a set of actions. The categories of question differ in the integrand, and the integrands draw on five computations: the posterior, the trend, the evidence, the preposterior and the expected consequence.

### 2.3.1 The probability of an event as an integral over the state

Suppose that inference has produced the posterior distribution of the state,

$$p(\mathbf{z} \mid \mathbf{y}, M),$$

and that a question asks whether the state lies in some region $\mathcal{A}$ of the state space. For the hall, $\mathcal{A}$ might be the set of all combinations of temperature and log-conductance that satisfy a protected limit. The probability of the region is the posterior integrated over it,

$$\Pr(\mathbf{z} \in \mathcal{A} \mid \mathbf{y}, M) = \int_{\mathcal{A}} p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}, \tag{2.6}$$

which is equation (2.5) with a vector in place of one quantity. The same expression is written with the indicator function of the event,

$$\mathbf{1}_A(\mathbf{z}) = \begin{cases} 1 & \text{if event } A \text{ holds at } \mathbf{z}, \\ 0 & \text{otherwise,} \end{cases}$$

which extends the integral to the whole state space and lets the indicator select the region:

$$\Pr(A \mid \mathbf{y}, M) = \int \mathbf{1}_A(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z} = \mathbb{E}\big[\mathbf{1}_A(\mathbf{z}) \mid \mathbf{y}, M\big]. \tag{2.7}$$

The last form reads the probability of an event as the posterior expectation of its indicator. The equation states the central rule of the chapter: *the probability of an event is the posterior integrated over the states at which the event is true.* Many control calculations are variations of it, differing only in the event and in how the integral is evaluated.

### 2.3.2 Future events and the posterior predictive

Some events remain random even when the state is known exactly. A loop whose trip rate is known still trips at a random time; knowing the rate fixes only the probability that a run passes without a trip. For such a future event $A_{\text{future}}$, the model supplies the conditional probability $\Pr(A_{\text{future}} \mid \mathbf{z}, M)$, and the question asks for its probability given the evidence. The derivation uses three rules. By the law of total probability, the probability of the event is the integral of its joint probability with the state:

$$\Pr(A_{\text{future}} \mid \mathbf{y}, M) = \int \Pr(A_{\text{future}}, \mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}.$$

By the product rule, the joint factors into a conditional and a marginal:

$$\Pr(A_{\text{future}} \mid \mathbf{y}, M) = \int \Pr(A_{\text{future}} \mid \mathbf{z}, \mathbf{y}, M)\, p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}.$$

Finally, the model assumes that the future event depends on the past evidence only through the state: once $\mathbf{z}$ is known, the readings and records of the past add nothing about it. Under that conditional independence, $\Pr(A_{\text{future}} \mid \mathbf{z}, \mathbf{y}, M) = \Pr(A_{\text{future}} \mid \mathbf{z}, M)$, and

$$\Pr(A_{\text{future}} \mid \mathbf{y}, M) = \int \Pr(A_{\text{future}} \mid \mathbf{z}, M)\, p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}. \tag{2.8}$$

This is the posterior predictive probability, the foundation of every prediction of a future trip, reading or excursion. Equation (2.7) is the special case in which the conditional probability is an indicator, zero or one at each state. The difference between the two matters as soon as the event is a future trip rather than a property of the present plant, which Section 2.4 shows with the loop's trip log.

### 2.3.3 The general forms of the five categories

Control questions about a plant are many, but the mathematical objects that answer them are few, and the book classifies questions by the object rather than by the word on the commissioning sheet. Five categories cover every question the book asks, each with its general forms, written here in the notation of Table 2.2: $p(\mathbf{z} \mid \mathbf{y})$ is the posterior density of the state; $f(\mathbf{z})$ is a quantity derived from the state, an entry being the case $f(\mathbf{z}) = z_i$, and $f^*$ a value of it, a requirement in a threshold question; $y_{\text{new}}$ is a future observation; $M_j$ is a complete model of the plant, a drawing with its rows; and $\mathbf{\Sigma}_a$ is the covariance of the state after a candidate measurement $a$. Nothing of any particular plant appears in the forms. Two further forms, of a trend (Section 2.3.4) and of an expected consequence (Section 2.3.5), belong to computations that the categories call on rather than to a category of their own. Every later chapter opens its derivation with one of these forms and then specializes it to its plant.

**Estimation.** An estimation question asks what a quantity is and how well it is known. Its general form is the posterior distribution of the quantity, the probability that it lies at or below each value $f^*$,

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \tag{2.9}$$

which is equation (2.7) for the event $f(\mathbf{z}) \le f^*$, read as a function of $f^*$. It is summarized by a named location, the mean, the median or the mode, and a spread; the mean and the variance are

$$\begin{aligned} \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] &= \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \\ \mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y}) &= \int \big(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]\big)^2\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \end{aligned} \tag{2.10}$$

the median is the value $f^*$ at which equation (2.9) equals $\frac{1}{2}$, and the mode is the value at which the density of $f(\mathbf{z})$ is largest. The uncertainty is also stated as the central 95 % credible interval $[f_L, f_U]$, whose endpoints leave 2.5 % of posterior probability in each tail:

$$\Pr(f(\mathbf{z}) \le f_L \mid \mathbf{y}) = 0.025, \qquad \Pr(f(\mathbf{z}) \le f_U \mid \mathbf{y}) = 0.975. \tag{2.11}$$

The quantiles of a monotone function of a quantity are the images of that quantity's quantiles, in the same order for an increasing function and in reverse order for a decreasing one; the median is a quantile and maps the same way, while the mean and the mode in general do not. Under the Laplace posterior $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$ an entry is read off directly, with location $\hat{z}_i$ and spread $\sqrt{\Sigma_{ii}}$, and a derived quantity through its gradient $\mathbf{g} = \partial f / \partial \mathbf{z}$ at $\hat{\mathbf{z}}$, with location $f(\hat{\mathbf{z}})$ and spread $\sqrt{\mathbf{g}^\top \mathbf{\Sigma} \mathbf{g}}$ by the delta method. The form is the posterior seen through one quantity: a distribution, of which any single number reported is a named summary.

**Prediction under a threshold.** A prediction under a threshold asks whether the plant meets a requirement. The requirement is written as the event $f(\mathbf{z}) \geq f^*$, "the requirement holds", with $f^*$ the required value, and its general form is the posterior probability of that event, compared with an assurance level $\kappa$, the posterior probability the requirement must reach:

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \geq f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad \text{accept if } \Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) \geq \kappa. \quad (2.12)$$

A requirement that a quantity stay below an upper limit, such as a temperature $T \leq T_{\max}$, is the same form with $f(\mathbf{z}) = -T$ and $f^* = -T_{\max}$; equivalently, the probability of the requirement is one minus the probability of the complementary event, the exceedance $T > T_{\max}$, and the rule is applied to the requirement, never to the exceedance. A joint event is the same integral with the indicator of an intersection. The required value is the limit on the protected quantity and the assurance level is the plant's policy, and the form keeps the two apart; a limit held with a stated probability is what a control engineer calls a chance constraint, with $\kappa = 1 - \alpha$ for a stated risk $\alpha$. When the event lies in the future and stays random even when the state is known, the tail is taken of the predictive distribution of the future observation,

$$p(y_{\text{new}} \mid \mathbf{y}) = \int p(y_{\text{new}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \quad (2.13)$$

of which equation (2.8) is the case of an event, with $y_{\text{new}}$ the indicator of $A_{\text{future}}$. The predictive is the randomness the model gives the future outcome, averaged over what is not known about the state. When the quantity moves in time and the requirement concerns a date, the event is a crossing, and its form is the trend form of Section 2.3.4.

**Hypothesis comparison.** A hypothesis comparison asks which of two descriptions of the plant is true. Its general form is the ratio of the evidences of two complete models fitted to the same evidence, the Bayes factor,

$$B_{AB} = \frac{p(\mathbf{y} \mid M_A)}{p(\mathbf{y} \mid M_B)}, \quad (2.14)$$

where the evidence $p(\mathbf{y} \mid M)$ of a model is the probability of the observed evidence with the model's own state integrated out over its prior. The ratio multiplies the prior odds of the two models into their posterior odds, and because each evidence integrates its own state out, it compares the drawings, not the values of their entries. The Bayes factor $B_{10}$ of the notation table is the case $M_A = M_1$ against $M_B = M_0$.

**Diagnosis.** A diagnosis asks where and what a fault is. Its general form is hypothesis comparison run over a family $M_1, \ldots, M_J$ of candidate faults generated from the plant graph, one for each component or reading that could be at fault, with the answer returned as the ranking of

$$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\, \Pr(M_j), \qquad j = 1, \ldots, J. \quad (2.15)$$

The answer is the order of the candidates with the posterior probability of each, not a single verdict.

**Design.** A design question asks what to measure or do next, before the measurement exists. Its general form is the preposterior expectation: for a candidate measurement $a$ whose outcome $\tilde{y}_a$ is not yet known, the expected value, over the outcomes weighted by their predictive distribution, of a quantity $u$ of the posterior that the outcome would produce,

$$U(a) = \mathbb{E}_{\tilde{y}_a \mid \mathbf{y}}\Big[\, u\big(p(\mathbf{z} \mid \mathbf{y}, \tilde{y}_a)\big)\Big] = \int u\big(p(\mathbf{z} \mid \mathbf{y}, \tilde{y}_a)\big)\, p(\tilde{y}_a \mid \mathbf{y})\, \mathrm{d}\tilde{y}_a, \quad (2.16)$$

where $p(\tilde{y}_a \mid \mathbf{y})$ is the predictive (2.13) of the outcome; for an outcome that takes discrete values, such as a number of trips in a trial period, the integral is a sum over those values weighted by their predictive probabilities. The quantity $u$ is one of four: the spread of a derived quantity after the outcome; the information gained; the indicator that an acceptance rule (2.12) is met after the outcome, which makes $U(a)$ the probability that the trial ends with the requirement accepted; or the expected loss of the best action once the outcome is known. The candidates are ranked by $U(a)$, the largest first when $u$ is an information gain or an acceptance indicator, the smallest first when $u$ is a spread or a loss, and a candidate's cost enters as a loss of its own. When the posterior is Gaussian and the candidate is a linear Gaussian row, $r_a(\mathbf{z}) = (\mathbf{h}_a\mathbf{z} - \tilde{y}_a)/\sigma_a$ with $\mathbf{h}_a$ the row vector of its sensitivities and $\sigma_a$ the standard deviation of its noise, the covariance of the updated posterior does not depend on the outcome, and for a spread or an information gain the form needs only two objects: the preposterior covariance,

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Lambda}_a\right)^{-1}, \qquad \boldsymbol{\Lambda}_a = \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a = \frac{\mathbf{h}_a^{\mathsf{T}}\mathbf{h}_a}{\sigma_a^2}, \tag{2.17}$$

with $\mathbf{J}_a = \mathbf{h}_a/\sigma_a$ the Jacobian of the candidate row, and the expected information gain, in nats,

$$\mathrm{EIG}(a) = \tfrac{1}{2}\ln\frac{\det\boldsymbol{\Sigma}}{\det\boldsymbol{\Sigma}_a}. \tag{2.18}$$

The spread of a derived quantity $f$ after the measurement is then $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_a\mathbf{g}}$, with the same gradient $\mathbf{g}$ at $\hat{\mathbf{z}}$. For a count, demand or lifetime row the candidate's information is the expected information: $\boldsymbol{\Lambda}_a = \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a$ evaluated at $\hat{\mathbf{z}}$ with the expected counts of the trial in place of observed ones, a count row entering $\mathbf{J}_a$ through the line $\sqrt{\mu_k}\,\partial \ln \mu_k/\partial\mathbf{z}$ of its working form (Section 4.2.7); equation (2.17) is then the Laplace preposterior, exact only to the order of the Laplace approximation, and it still does not depend on the outcome. A design question values a posterior that does not exist yet, through what the reading would do to it. When the candidate is an action that brings no new evidence, retuning a loop or changing a set point, there is no outcome to average over, the posterior after the action is the present one, and the form reduces to the expected consequence of Section 2.3.5.

### 2.3.4 The general form of a trend

Some questions concern a quantity that moves in time: whether a trip rate or a coil's derate is rising, when it will cross its limit, what it will be at the end of the run. By its integrand each is an estimation or a prediction under a threshold; what sets it apart is that the quantity is indexed by time, so that the state must carry how the quantity changes as well as where it stands. Write the quantity, on the scale on which the state carries it (a log-rate, a log-derate, a temperature), as $f_t(\mathbf{z})$, its value at time $t$. In the simplest drawing it is a straight line in time about the center $\bar{t}$ of the record, with a level entry $\alpha$, the quantity at $\bar{t}$, and a *trend entry* $\gamma$, its change per unit time:

$$\mathbf{z} = \begin{bmatrix}\alpha\\ \gamma\end{bmatrix}, \qquad f_t(\mathbf{z}) = \alpha + \gamma\,(t - \bar{t}) = \mathbf{g}(t)^{\mathsf{T}}\mathbf{z}, \qquad \mathbf{g}(t) = \frac{\partial f_t}{\partial \mathbf{z}} = \begin{bmatrix}1\\ t - \bar{t}\end{bmatrix}. \tag{2.19}$$

The evidence arrives in windows, one row per window. Window $k$, centered at the time $t_k$, gives the row $r_k(\mathbf{z})$, which sees the state only through the quantity at its own time, $f_{t_k}(\mathbf{z})$, so its Jacobian row is the gradient $\mathbf{g}(t_k)^{\mathsf{T}}$ times the row's sensitivity $d_k = \partial r_k/\partial f_{t_k}$: $1/\sigma_k$ for a reading row of standard deviation $\sigma_k$, and $\sqrt{\mu_k}$, the square root of the window's expected count,

for a count row in the working form of Section 4.2.7. The rows of the record stack to

$$\mathbf{J} = \mathbf{D}\,\mathbf{X}, \qquad \mathbf{X} = \begin{bmatrix} \mathbf{g}(t_1)^{\mathsf{T}} \\ \mathbf{g}(t_2)^{\mathsf{T}} \\ \vdots \end{bmatrix}, \qquad \mathbf{D} = \operatorname{diag}(d_1, d_2, \ldots),$$
$$\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J} = \begin{bmatrix} \sum_k d_k^2 & \sum_k d_k^2\,(t_k - \bar{t}) \\ \sum_k d_k^2\,(t_k - \bar{t}) & \sum_k d_k^2\,(t_k - \bar{t})^2 \end{bmatrix}, \qquad \mathbf{\Sigma} = \mathbf{\Lambda}^{-1}. \tag{2.20}$$

The trend entry of the information, $\sum_k d_k^2 (t_k - \bar{t})^2$, weighs each window's distance from the center squared, while the windows and their precisions $d_k^2$ add linearly: widening the span of a record raises the information on the trend faster than adding windows inside it.

The posterior of this state is read along time in three ways, and the three make up the general form of a trend:

$$\begin{aligned} \Pr(\gamma > 0 \mid \mathbf{y}) &= \int \mathbf{1}[\gamma > 0]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \\ \Pr(t^* \le s \mid \mathbf{y}) &= \int \mathbf{1}[\gamma > 0]\, \mathbf{1}\big[f_s(\mathbf{z}) \ge f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \\ \Pr(f_t(\mathbf{z}) \le f^* \mid \mathbf{y}) &= \int \mathbf{1}\big[f_t(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad t \text{ beyond the record.} \end{aligned} \tag{2.21}$$

The first line is the sign of the trend, the probability that the quantity is rising: the threshold form on the trend entry. The second is the distribution function of the crossing time $t^*(\mathbf{z})$, the time at which the quantity reaches the level $f^*$ from below. A rising path, $\gamma > 0$, has crossed by a time $s$ if and only if it is at or above $f^*$ at $s$, whether $s$ lies before or after $\bar{t}$, and a path with $\gamma \le 0$ never crosses, a mass $\Pr(\gamma \le 0 \mid \mathbf{y})$ at $t^* = \infty$: the joint indicator is the crossing, the first passage of a rising path. The threshold probability $\Pr(f_s(\mathbf{z}) \ge f^* \mid \mathbf{y})$ alone is not the crossing, because it also counts the states with $\gamma \le 0$ that stand at or above the level at $s$, paths that started above it and have never crossed it; the two agree only when the posterior puts no such state there. The third is the forecast, the estimation form at a time beyond the last window of the record. Nothing of a particular plant appears in the three lines; a plant enters through the scale of $f_t$, the rows of its windows and the level $f^*$.

Under the Laplace posterior $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$ the quantity at any time is linear in the state, so the sign and the forecast are exact normal tails in two numbers, the center $\mathbf{g}(t)^{\mathsf{T}}\hat{\mathbf{z}}$ and the variance $\mathbf{g}(t)^{\mathsf{T}}\mathbf{\Sigma}\,\mathbf{g}(t)$, and the crossing is an exact bivariate normal probability of the pair $(\gamma, f_s)$, whose correlation $\varrho(s)$ is read from the same $\mathbf{\Sigma}$:

$$\begin{aligned} &\Pr(\gamma > 0 \mid \mathbf{y}) \approx \Phi\!\left(\frac{\hat{\gamma}}{\sqrt{\Sigma_{\gamma\gamma}}}\right), \qquad \Pr(t^* \le s \mid \mathbf{y}) \approx \Phi_2\!\left(\frac{\hat{\gamma}}{\sqrt{\Sigma_{\gamma\gamma}}},\ \frac{\mathbf{g}(s)^{\mathsf{T}}\hat{\mathbf{z}} - f^*}{\sqrt{\mathbf{g}(s)^{\mathsf{T}}\mathbf{\Sigma}\,\mathbf{g}(s)}};\ \varrho(s)\right), \\ &\mathbf{g}(t)^{\mathsf{T}}\mathbf{\Sigma}\,\mathbf{g}(t) = \Sigma_{\alpha\alpha} + 2\,(t - \bar{t})\,\Sigma_{\alpha\gamma} + (t - \bar{t})^2\,\Sigma_{\gamma\gamma}, \qquad \varrho(s) = \frac{\Sigma_{\alpha\gamma} + (s - \bar{t})\,\Sigma_{\gamma\gamma}}{\sqrt{\Sigma_{\gamma\gamma}\,\mathbf{g}(s)^{\mathsf{T}}\mathbf{\Sigma}\,\mathbf{g}(s)}}, \end{aligned} \tag{2.22}$$

where $\Phi_2(u_1, u_2; \varrho)$ is the distribution function of two standard normals of correlation $\varrho$, the probability that the first lies below $u_1$ and the second below $u_2$. It is one integral, $\Phi_2(u_1, u_2; \varrho) = \int_{-\infty}^{u_1} \varphi(x)\, \Phi\big((u_2 - \varrho x)/\sqrt{1 - \varrho^2}\big)\, \mathrm{d}x$, with $\varphi$ the standard normal density, and it lies below the single tail $\Phi(u_2)$, the *threshold line*, by the mass of the states that stand beyond the level without a rising trend. The sign is the single tail with the gradient $(0,\ 1)^{\mathsf{T}}$, which reads the trend entry alone, so that its variance is $\Sigma_{\gamma\gamma}$; the forecast is the same with the time asked. The variance of the quantity grows with the square of the distance $t - \bar{t}$, the lever arm: a forecast pays for the uncertainty of the trend with that square. The crossing time itself, $t^* = \bar{t} + (f^* - \alpha)/\gamma$ for $\gamma > 0$, is a ratio with the trend entry in its denominator. When the posterior of $\gamma$ has a

positive density at zero, as a Gaussian posterior always has, $t^*$ has no posterior mean and no variance, so it is reported by its median, the date at which the second line is one half, which lies near the plug-in date $t^*(\hat{\mathbf{z}})$ when little of the posterior stands beyond the level without a rising trend, by the band whose ends solve the second line at 0.025 and 0.975, and by the mass that never crosses; the delta band $t^*(\hat{\mathbf{z}}) \pm 1.96\sqrt{\mathbf{g}_*^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}_*}$, with $\mathbf{g}_* = \partial t^*/\partial \mathbf{z}$ at $\hat{\mathbf{z}}$, is a finite number standing in for an infinite one. For the trip log of a sister loop's older valve, counted in six periods of forty days and rising from 5 to 14 trips a period (Exercise 2.5), the band of the crossing of the site's valve-replacement limit of 0.5 trips a day runs from 5.90 to 28.3 periods by inversion, with period $j$ centered at $t = j$ and the log ending at $t = 6.5$, against a delta band from 4.44 to 11.45 periods that is wrong at both ends. A flat or falling line on that log stands at or above the limit at those dates with a mass of at most $5.1 \times 10^{-11}$, so the first passage and the threshold line give the same median and band to every digit printed, and the band is read from the line. Chapter 19 reads the same form on the load and weather along a plan.

### 2.3.5 The general form of an expected consequence

Other questions ask what to do: whether to retune the loop before the night peak, whether to raise a set point, which loop to take to manual. Each compares actions $a$ from a set of candidates by what they are expected to cost. Let $L(a, \mathbf{z})$ be the loss of taking action $a$ when the state is $\mathbf{z}$, its consequence in whatever unit the decision is taken in: reais, minutes above the limit, energy. It is a derived quantity of the state, one for each action. The general form of an expected consequence is its posterior expectation, the expected loss $\mathcal{R}(a)$ of the action, and the action that minimizes it. When the action costs $c_a$ for certain and leaves events $e$ that occur with probability $\Pr(e \mid \mathbf{z}, a)$ and have the consequence $C_e$ each, the loss given the state is the cost of the action plus the consequence of each event weighted by its probability, and by the predictive (2.8) the posterior expectation turns each conditional probability into its predictive one:

$$\begin{aligned} \mathcal{R}(a) &= \mathbb{E}\big[L(a,\mathbf{z}) \mid \mathbf{y}\big] = \int L(a,\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad a^* = \arg\min_a \mathcal{R}(a), \\ L(a,\mathbf{z}) &= c_a + \sum_e C_e \Pr(e \mid \mathbf{z}, a) \quad \Longrightarrow \quad \mathcal{R}(a) = c_a + \sum_e C_e \Pr(e \mid \mathbf{y}, a). \end{aligned} \tag{2.23}$$

The second line is the rule that risk is probability times consequence, summed over the events and added to the certain cost of the action. Nothing of a particular plant appears in it either; a plant enters through the events, the probabilities its posterior gives them and the consequence schedule $c_a$, $C_e$ that the plant's owner supplies.

Under the Laplace posterior the expected consequence is read from the same objects as every other answer. With $\mathbf{g}_a = \partial L(a,\mathbf{z})/\partial \mathbf{z}$ and $\mathbf{H}_a = \partial^2 L(a,\mathbf{z})/\partial \mathbf{z}\,\partial \mathbf{z}^{\mathsf{T}}$ at $\hat{\mathbf{z}}$, the expansion of the loss to second order about $\hat{\mathbf{z}}$, averaged over $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, keeps the constant and the curvature term, since $\mathbb{E}[\mathbf{z} - \hat{\mathbf{z}} \mid \mathbf{y}] = \mathbf{0}$ and $\mathbb{E}[(\mathbf{z} - \hat{\mathbf{z}})^{\mathsf{T}}\mathbf{H}_a(\mathbf{z} - \hat{\mathbf{z}}) \mid \mathbf{y}] = \operatorname{tr}(\mathbf{H}_a\boldsymbol{\Sigma})$; the spread follows from the gradient, and the comparison of two actions from the gradient of their difference:

$$\begin{aligned} \mathcal{R}(a) &\approx L(a,\hat{\mathbf{z}}) + \tfrac{1}{2}\operatorname{tr}(\mathbf{H}_a\boldsymbol{\Sigma}), \qquad \operatorname{sd}\big(L(a,\mathbf{z}) \mid \mathbf{y}\big) \approx \sqrt{\mathbf{g}_a^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}_a}, \\ \Pr\big(L(a,\mathbf{z}) < L(a',\mathbf{z}) \mid \mathbf{y}\big) &\approx \Phi\left(\frac{\mathcal{R}(a') - \mathcal{R}(a)}{\sqrt{(\mathbf{g}_a - \mathbf{g}_{a'})^{\mathsf{T}}\boldsymbol{\Sigma}\,(\mathbf{g}_a - \mathbf{g}_{a'})}}\right). \end{aligned} \tag{2.24}$$

The loss at the most probable state, $L(a,\hat{\mathbf{z}})$, misses the curvature term. The curvature is taken in the coordinates of the state, and its sign there, not in a physical quantity, sets the direction of the second-order term. The trip log of Section 2.4.4 shows the difference over the next 100 h of operation, about eight night runs. The probability of no trip, $\mathrm{e}^{-\lambda\tau}$, is convex in the rate $\lambda$, so by Jensen's inequality the predictive 0.7725 lies above the plug-in 0.7659; but the state carries the log-rate $\omega$, in which the probability is concave, with curvature $-0.150$ at $\hat{\omega}$, and

the second-order reading with $\sigma_\omega^2 = 0.25$ moves the plug-in down to 0.747, close to the average 0.748 over the Laplace normal of the log-rate and away from the exact 0.7725. The term is right about the Laplace posterior and wrong about the trips, because the error lies in the Laplace posterior itself; where the second-order term and an exact route move the plug-in in opposite directions, the value read from the state is stated with the exact one beside it, and the gap is what the Laplace posterior costs. The second line of equation (2.24) asks how firm a ranking is: the probability that action $a$ costs less than $a'$ at the true state. The variance of the difference is the sum of the two variances minus $2\,\mathbf{g}_a^\mathsf{T}\boldsymbol{\Sigma}\,\mathbf{g}_{a'}$, so the band of the comparison is narrower than the two bands taken apart exactly when $\mathbf{g}_a^\mathsf{T}\boldsymbol{\Sigma}\,\mathbf{g}_{a'} > 0$, when the two losses move together across the posterior. Section 2.5.4 applies the form to the choice between retuning the loop before the night peak and leaving the tuning.

### 2.3.6 The five computations beneath the forms

Five computations sit underneath the forms, each defined by the object it returns: the posterior of the state, the trend, the evidence, the preposterior and the expected consequence.

**The posterior.** The posterior of the state follows from Bayes' rule, the likelihood of the evidence times the prior divided by their integral over the state. For Gaussian rows divided by their standard deviations it is the exponential of minus the objective of the rows,

$$p(\mathbf{z} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \mathbf{z})\,p(\mathbf{z})}{p(\mathbf{y})} \propto \exp\bigl[-\mathcal{L}(\mathbf{z})\bigr], \qquad \mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,r(\mathbf{z})^\mathsf{T} r(\mathbf{z}), \tag{2.25}$$

and a count, demand or lifetime row puts its negative log-probability term into $\mathcal{L}$ instead. Under the Laplace approximation the posterior is the normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ centered at the MAP state $\hat{\mathbf{z}}$, the minimum of $\mathcal{L}$, with the covariance $\boldsymbol{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ built from the Jacobian $\mathbf{J}$ of the rows; Chapter 3 derives the normal equations and the covariance, and Chapter 4 the Laplace approximation. The estimation, threshold and predictive forms are integrals against this posterior, read for an entry $\hat{z}_i$ with $\sqrt{\Sigma_{ii}}$ or for a derived quantity through $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$.

**The trend.** The trend is the posterior over a state indexed by time or by window, with a trend entry among its entries, read along time: as the probability that the trend entry is positive, as a crossing time with its band, or as a forecast, the three lines of equation (2.21). Its building blocks are those of the posterior, $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ from the rows of the windows (2.20), with the gradient $\mathbf{g}(t)$ of the quantity at the time asked. It is a reading of the posterior, and it is named as a computation of its own because its object, its check and its failures are its own: the object is a function of the time asked rather than one number; the check orders the record's residuals in time and inverts the second line of that form for the band of the crossing instead of linearizing the crossing time, a band that is the threshold line's band only where the first passage and the threshold line agree; and it fails where no other reading does, in a forecast that pays for the uncertainty of the trend with the square of its lever arm, in a crossing time without a mean when the posterior of the trend reaches zero, and in the shape of the line, which the record tests only inside its own span.

**The evidence.** The denominator of equation (2.25), written with its model explicit, is the evidence of the model, the probability of the observed evidence with the model's own unknowns integrated out over their prior:

$$p(\mathbf{y} \mid M_j) = \int p(\mathbf{y} \mid \mathbf{z}, M_j)\,p(\mathbf{z} \mid M_j)\,\mathrm{d}\mathbf{z}. \tag{2.26}$$

It is the normalizing constant of the posterior within model $M_j$, and it is the only place where the comparison between models enters: the posterior inside either model, however sharp, does

not say which model is true. Under the Laplace approximation it is read from the model's own fit, as the volume of its posterior times the height of the integrand at its MAP state, $p(\mathbf{y} \mid M_j) \approx (2\pi)^{d_j/2} \det(\boldsymbol{\Sigma}_j)^{1/2}\, p(\mathbf{y} \mid \hat{\mathbf{z}}_j, M_j)\, p(\hat{\mathbf{z}}_j \mid M_j)$ with $d_j$ the number of entries of its state, exact when the objective is quadratic (Chapter 4). The comparison and diagnosis forms are a ratio and a ranking of it.

**The preposterior.** The preposterior is the expected posterior under a measurement not yet taken: the posterior that each outcome $\tilde{y}_a$ would produce, averaged over the predictive distribution of the outcomes, the inner object of the design form (2.16). For a linear Gaussian candidate row its covariance does not depend on the outcome, and equation (2.17) follows from the rows. Let the current posterior of the state be approximately Gaussian with mean $\hat{\mathbf{z}}$ and covariance $\boldsymbol{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$, and let the proposed measurement depend on the state through its sensitivities, $\tilde{y}_a = \mathbf{h}_a\mathbf{z} + \varepsilon$ to first order, with independent Gaussian noise $\varepsilon$ of variance $\sigma_a^2$. Divided by its standard deviation, the candidate row is $r_a(\mathbf{z}) = (\mathbf{h}_a\mathbf{z} - \tilde{y}_a)/\sigma_a$, and its Jacobian is the row vector $\mathbf{J}_a = \partial r_a/\partial \mathbf{z} = \mathbf{h}_a/\sigma_a$. The objective after the measurement is the current one plus that row,

$$\mathcal{L}_a(\mathbf{z}) = \tfrac{1}{2}\, r(\mathbf{z})^\mathsf{T} r(\mathbf{z}) + \tfrac{1}{2}\, r_a(\mathbf{z})^2,$$

and its curvature, the information, is the curvature of the stacked Jacobian $[\mathbf{J};\ \mathbf{J}_a]$, the sum $\mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{J}_a^\mathsf{T}\mathbf{J}_a$: the information of independent rows adds. With $\mathbf{J}^\mathsf{T}\mathbf{J} = \boldsymbol{\Sigma}^{-1}$, the covariance after the measurement under the linear-Gaussian approximation is the inverse of that sum, which is equation (2.17). When the current rows include exact constraints, $\boldsymbol{\Sigma}$ is singular in the limit $\sigma_p \to 0$ and $\boldsymbol{\Sigma}^{-1}$ does not exist there. The limit is then taken through the equivalent update form, which uses $\boldsymbol{\Sigma}$ alone,

$$\boldsymbol{\Sigma}_a = \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{J}_a^\mathsf{T}\big(\mathbf{I} + \mathbf{J}_a\boldsymbol{\Sigma}\mathbf{J}_a^\mathsf{T}\big)^{-1}\mathbf{J}_a\boldsymbol{\Sigma}, \tag{2.27}$$

where $\mathbf{I}$ is the identity of the size of the number of candidate rows. The value $\tilde{y}_a$ of the measurement does not appear in either form, only its Jacobian, its sensitivity divided by its noise. That is why equation (2.17) predicts how the uncertainty would shrink before the measurement is taken, and why it is one foundation of sensor placement and experimental design. The information gain (2.18) is the standard result for a Gaussian prior and a linear Gaussian measurement; where both determinants vanish in the limit of exact rows, the determinant identity $\det\boldsymbol{\Sigma}/\det\boldsymbol{\Sigma}_a = \det(\mathbf{I} + \mathbf{J}_a\boldsymbol{\Sigma}\mathbf{J}_a^\mathsf{T})$, which also uses $\boldsymbol{\Sigma}$ alone, evaluates it.

**The expected consequence.** This computation returns the expected loss of each candidate action over the posterior, $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$, and the choice of the action that minimizes it, equation (2.23); summed over events, it is probability times consequence. Its building blocks are $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ with the gradient $\mathbf{g}_a$ and the curvature $\mathbf{H}_a$ of each action's loss, equation (2.24). It is the expectation of one more derived quantity, and it is named as a computation of its own because its object is a choice among actions rather than a description of the state, because its check is the curvature term that separates it from the loss at the most probable state together with the probability that the ranking holds, and because its failures are its own: a second-order reading whose sign follows the coordinates of the state, a ranking that is clear in expectation and uncertain at the true state, and a consequence schedule that the posterior does not contain, whose own uncertainty no band of the state carries.

Table 2.3 sets the five categories against the five computations: its first column is the category, its second the computation with the kind of question in the category that calls on it, and its third the object the computation returns. A category calls on more than one computation when its questions differ in kind.

In this book the posterior is used in every chapter that fits a state; the trend leads in Chapter 19; the evidence in Chapters 4 and 27; the preposterior at step 7 of every chapter

Table 2.3: The five categories of question, the computations that answer them, and the object each computation returns.

| Category | Computation | Object returned |
|---|---|---|
| estimation | posterior | an entry, $\hat{z}_i$ with $\sqrt{\Sigma_{ii}}$; a derived quantity, $f(\hat{\mathbf{z}})$ with $\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}$ |
| | trend, for the quantity at a stated time and for the distribution of a crossing time | $\mathbf{g}(t)^\mathsf{T}\hat{\mathbf{z}}$ with $\mathbf{g}(t)^\mathsf{T}\boldsymbol{\Sigma}\,\mathbf{g}(t)$; the median crossing time and its band by inversion |
| | expected consequence, for the expected loss of one course of action | $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$ |
| prediction under a threshold | posterior | a tail of an entry or of a derived quantity; a predictive |
| | trend, for the sign of a trend and for a crossing by a date | $\Pr(\gamma > 0 \mid \mathbf{y})$; $\Pr(t^* \le s \mid \mathbf{y})$ |
| hypothesis comparison | evidence | a Bayes factor $B_{AB} = p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$ |
| diagnosis | evidence | the ranking of $\Pr(M_j \mid \mathbf{y})$ over a family generated from the graph |
| design | preposterior, for the next measurement | $\boldsymbol{\Sigma}_a$, the information gain, the value of information |
| | expected consequence, for the choice of an action | $a^* = \arg\min_a \mathcal{R}(a)$; the probability that the ranking holds |

that walks the steps, and as the lead computation in Chapters 16, 22 and 23; and the expected consequence in Section 2.5.4 and in Chapters 7, 12, 13, 14, 17, 21, 22 and 28.

### 2.3.7 The finer question types inside the five categories

A finer list of question types is useful in practice, because each type has a characteristic wording. Table 2.4 lists twelve of them. The first column is the type, the second an example in the engineer's words, the third the mathematical target, and the last two the category of the book it belongs to and the computation beneath it. The list does not limit the applications; it collects the mathematical patterns that recur across them.

Most rows fall into a category directly; five need a sentence. A *joint event* such as "both loops run clean" is an event like any other, a region of the state space, so its probability is equation (2.7) with the indicator of an intersection; the wording "both" changes the region, not the category. A *trend* question asks for threshold events, on the trend entry, on the quantity at a date and, for a crossing, on the two jointly, so its category is prediction under a threshold; its computation is the trend, because the state is indexed by time and the crossing time is a ratio whose distribution function must be inverted to bound it. An *expectation* of a consequence, $\mathbb{E}[L \mid \mathbf{y}] = \int L(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$, is the mean of the marginal of the derived quantity $L(\mathbf{z})$: its category is estimation, applied to a loss, and its computation is the expected consequence, whose curvature term separates it from the loss at the most probable state; its tail $\Pr(L > l \mid \mathbf{y})$ would be a prediction under a threshold. An *importance* measure (in control, a gain or a sensitivity) is a derivative of a derived quantity with respect to an entry of the state, the gradient that the delta method already uses, and it belongs with estimation. A *decision* among actions is a design question answered by the expected consequence, with the preposterior as well when a reading comes before the action.

Table 2.4: Twelve question types, their mathematical targets, and the category and computation each belongs to.

| Question type | Example | Mathematical target | Category | Computation |
|---|---|---|---|---|
| Estimation | What is the loop's trip rate? | $p(\lambda \mid \mathbf{y})$ | estimation | posterior |
| Threshold probability | Will the hall stay within its limit? | $\Pr(T \le T_{\max} \mid \mathbf{y})$ | prediction under a threshold | posterior |
| Joint and conditional events | Will both loops run clean through the night? | $\Pr(\bar{A}_A \cap \bar{A}_B \mid \mathbf{y})$ | prediction under a threshold | posterior |
| Prediction | Will the night run pass without a trip? | $\Pr(T > 12\,\mathrm{h} \mid \mathbf{y})$ | prediction under a threshold | posterior |
| Trend and crossing | Is the trip rate rising, and when will it pass the valve-replacement limit? | $\Pr(\gamma > 0 \mid \mathbf{y})$, $\Pr(t^* \le s \mid \mathbf{y})$ | prediction under a threshold | trend |
| Derived quantities | What is the process gain to the load? | $p(K_Q(\mathbf{z}) \mid \mathbf{y})$ | estimation | posterior |
| Model comparison | Is the sensor off, or the path changed? | $\Pr(M_j \mid \mathbf{y})$ | hypothesis comparison | evidence |
| Diagnosis | Which element of the loop is at fault? | $\Pr(F_j \mid \mathbf{y})$ | diagnosis | evidence |
| Risk and expectation | What does an excursion tonight cost on average? | $\mathbb{E}[L \mid \mathbf{y}]$ | estimation | expected consequence |
| Importance | Which entry moves the hall temperature most? | $\partial T/\partial z_i$, an entry of the gradient $\mathbf{g}$ | estimation | posterior |
| Information and experiment design | Which step test should run next? | expected information gain or value of information | design | preposterior |
| Decision optimization | Retune before the peak, or leave the tuning? | $\arg\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$ | design | expected consequence, with the preposterior when a test comes first |

### 2.3.8 The control questions of the book

The book is organized around the questions that come up when a loop is commissioned and run, one leading each chapter from Part II on, grouped by the part of the book they lead. Table 2.5 lists the twenty-six of them in the control engineer's words. Its first column is the chapter that leads with the question; the second the question; the third the general form it takes, by equation number; the fourth its category; and the last the computation beneath it, or the two computations when the question calls on two.

Table 2.5: The control engineer's twenty-six questions, one per chapter from Chapter 5 to Chapter 30, with the general form, the category and the computation of each.

| Chapter | Control engineer's question | General form | Category | Computation |
|---|---|---|---|---|
| *The closed-loop problem* | | | | |
| 5 | What does closing the loop change in the graph? | (2.9), at a scenario node | estimation | posterior |
| 6 | What must stay inside its limit, with what probability? | (2.12) | prediction under a threshold | posterior |
| 7 | What is the problem every controller is solving part of? | (2.23), with the other forms inside it | design | expected consequence, with all five on one example |
| 8 | How likely is this tuning to be stable, given the band on the plant? | (2.12), on a derived quantity | prediction under a threshold | posterior |
| *The belief* | | | | |
| 9 | Is the filter a separate device? | (2.10), time-stacked | estimation | posterior |
| 10 | What does one new reading do to the state the loop acts on? | (2.10), one row added | estimation | posterior |
| 11 | Can the sensors resolve what the controller needs, and what does the loop do when they cannot? | (2.10), with a row removed | estimation | posterior |
| *Deriving the control law* | | | | |
| 12 | When does the band change the move at all, and what is the law when it does not? | (2.23) | design | expected consequence |
| 13 | Where does the integrator come from, and what can it not remove? | (2.10); (2.23) | estimation; design | posterior; expected consequence |
| 14 | What does a given tuning optimize, if anything? | (2.23) | design | expected consequence |
| 15 | What happens to the law and the band at the stop? | (2.9), at a scenario node | estimation | posterior |
| *Limits and the horizon* | | | | |
| 16 | What set point holds the inlets at 95 %, and what is a tighter posterior worth? | (2.12), at a scenario node; (2.16) | prediction under a threshold; design | posterior; preposterior |
| 17 | Where does cost put the set point, and which risk does it imply? | (2.23) | design | expected consequence |
| 18 | What is the plan over the next hour, and what does a fixed plan give up? | (2.12), at scenario nodes across time; the plan's objective deferred to that chapter | prediction under a threshold; design | posterior |
| 19 | How do the load and weather bands tighten the plan, and what does doing nothing risk? | (2.12), a joint tail; (2.21) | prediction under a threshold | posterior; trend |
| 20 | Why take only the first move, and what is a response worth? | (2.10), against the planned posterior | estimation | posterior |
| 21 | When should storage charge, given the band on price and load? | (2.23) | design | expected consequence |

| Chapter | Control engineer's question | General form | Category | Computation |
|---|---|---|---|---|
| *Learning while controlling* | | | | |
| 22 | Which move holds the limit and narrows the band, and when is that worth anything? | (2.23), (2.16) | design | expected consequence; preposterior |
| 23 | Which test signal identifies the plant fastest inside the constraint? | (2.16) | design | preposterior |
| 24 | Why does the loop hide the parameter, and which rows recover it? | (2.10), on a closed-loop record | estimation | posterior |
| 25 | How far does the set point move as the band shrinks over a season? | (2.12), at a scenario node, under a sequential posterior | prediction under a threshold | posterior |
| *Faults and reconfiguration* | | | | |
| 26 | Why does a well-tuned loop hide a fault, and what sees it anyway? | (2.13), (2.12) | prediction under a threshold | posterior |
| 27 | Which loop should switch, what is the new set point, and does it hold through the switch? | (2.15); (2.12), at a scenario node | diagnosis; prediction under a threshold | evidence; posterior |
| 28 | Which move holds under every ranked fault? | (2.12) under each fault of (2.15), against (2.23) | design | expected consequence |
| *Supervision* | | | | |
| 29 | When should the second unit start, and how does a cooling set point move an electrical constraint? | (2.12), on a margin and as a joint tail | prediction under a threshold | posterior |
| 30 | What do the three time scales share, and what does the operator see? | (2.10), three reads | estimation | posterior |

Most control questions are the posterior read at a pinned setting: eighteen of the twenty-six call on it, which is why one fit can serve a supervisory layer. The expected consequence, called on by eight, leads the statement of the problem in Chapter 7, and enters where the loss is not a limit, in Chapters 12, 14, 17 and 21; in Chapter 13 beside the posterior; in Chapter 22 beside the preposterior; and in Chapter 28, where it compares the moves that hold a limit under each ranked fault. The preposterior, called on by three, enters where a move also teaches, in Chapters 16, 22 and 23. The trend and the evidence are called on by one each.

## 2.4 Worked examples of estimation, prediction and joint events

Three pairs of questions are often confused on a commissioning sheet: the estimate of a quantity and the probability of a threshold event, the estimate of a trip rate and the prediction of a clean run, a joint event on two loops and the product of two marginal probabilities. In each pair the two questions are put to the same model and differ in the form put to it.

### 2.4.1 Estimation of the conductance of the hall

The engineer has measured the hall temperature at the present load of 30 kW and wants to know what the gain to the load is, which is to say what the conductance is (question A), and what will happen at the future load of $Q_2 = 60\,\mathrm{kW}$ (question B). Step 2 of the workflow has extended the graph with the scenario node of Figure 2.1, the hall at full load, whose temperature $T_2$ obeys the balance of equation (2.2),

$$T_2 = T_c + \frac{Q_2}{G}, \qquad T_c = 18^\circ\mathrm{C}, \quad Q_2 = 60\,\mathrm{kW}. \tag{2.28}$$

The conductance enters the state as the log-conductance $g = \ln(G/G_{\mathrm{ref}})$ with $G_{\mathrm{ref}} = 1\,\mathrm{kW/K}$ in this chapter's examples (Chapter 1 takes 6 kW/K; the reference only shifts $g$ by a constant).

For this example, suppose that a longer record of the hall has been summarized as one row on the log-conductance, the value $\hat{g}$ with the standard deviation $\sigma_g$,

$$r_g(\mathbf{z}) = \frac{g - \hat{g}}{\sigma_g}, \qquad \hat{g} = \ln 6.5 = 1.871802, \quad \sigma_g = 0.05. \tag{2.29}$$

These values are illustrative. The row stands for that longer record and is not the fit of Chapter 1, whose one reading of 0.5 K cannot give a spread of 5 %: there the MAP conductance is 5.917 kW/K and the probability of staying below 27.5°C at full load is 0.246, the same question under different evidence. How readings are reduced to a row of this kind is the subject of later chapters. The row is the only one that acts on $g$ alone, and the fit below returns the posterior of the third entry with center $\hat{g}$ and spread $\sqrt{\Sigma_{33}} = \sigma_g$.

Question A is an estimation question, of the general form (2.9). The quantity asked for is the conductance, a derived quantity of the log-conductance, $f(\mathbf{z}) = G = G_{\text{ref}}\,\mathrm{e}^{g}$, with $g$ the third entry of the state; it is increasing in that one entry. The answer is therefore read from three objects of the fit of the hall: the MAP state $\hat{\mathbf{z}}$, the Laplace covariance $\boldsymbol{\Sigma}$ and the gradient of $f$. The gradient has a single nonzero entry, $\mathbf{g} = (0,\ 0,\ G)^{\mathsf{T}}$, so the band $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} = G^2\,\Sigma_{33}$ uses only the third diagonal entry of $\boldsymbol{\Sigma}$; and because $f$ is increasing in one entry, its quantiles are the images of that entry's quantiles under the exponential. When the third entry is exactly normal, as it is here, $G$ is exactly lognormal, with median $G_{\text{ref}}\,\mathrm{e}^{\hat{z}_3}$, mean $G_{\text{ref}}\,\mathrm{e}^{\hat{z}_3+\Sigma_{33}/2}$ and central 95 % interval $G_{\text{ref}}\,\mathrm{e}^{\hat{z}_3 \pm 1.960\sqrt{\Sigma_{33}}}$, and these exact summaries are reported beside the Laplace band. The fit of the state of the extended graph is the building block.

The posterior of the whole state of the extended graph follows from the row (2.29) and the two balance rows as a fit. The state holds the hall temperature $T_1$ at the present load $Q_1 = 30\,\text{kW}$, the full-load temperature $T_2$ and the log-conductance,

$$\mathbf{z} = \begin{bmatrix} T_1 \\ T_2 \\ g \end{bmatrix}. \tag{2.30}$$

The rows are the energy balances of the two hall nodes, written in temperature form as in Chapter 1 and divided by a model tolerance $\sigma_p$ in kelvin, and the earlier inference, which enters as a row on $g$ of the same form as a prior row:

$$r(\mathbf{z}) = \begin{bmatrix} (T_1 - T_c - (Q_1/G_{\text{ref}})\,\mathrm{e}^{-g})/\sigma_p \\ (T_2 - T_c - (Q_2/G_{\text{ref}})\,\mathrm{e}^{-g})/\sigma_p \\ (g - \hat{g})/\sigma_g \end{bmatrix}, \qquad \mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\, r(\mathbf{z})^{\mathsf{T}} r(\mathbf{z}). \tag{2.31}$$

The exact balance of equation (2.28) is the limit $\sigma_p \to 0$. The Jacobian has one row per row of $r$ and one column per entry of $\mathbf{z}$, so it is $3 \times 3$; each entry is the sensitivity of a row to an entry of the state, in standard deviations of the row per unit of the entry:

$$\mathbf{J} = \frac{\partial r}{\partial \mathbf{z}} = \begin{bmatrix} 1/\sigma_p & 0 & (Q_1/G_{\text{ref}})\,\mathrm{e}^{-g}/\sigma_p \\ 0 & 1/\sigma_p & (Q_2/G_{\text{ref}})\,\mathrm{e}^{-g}/\sigma_p \\ 0 & 0 & 1/\sigma_g \end{bmatrix} = \begin{bmatrix} 100 & 0 & 461.5 \\ 0 & 100 & 923.1 \\ 0 & 0 & 20 \end{bmatrix}$$

at the most probable state with $\sigma_p = 0.01\,\text{K}$, in $\text{K}^{-1}$ in the first two columns. The third column says that a unit of $g$ moves the full-load balance twice as far as the present-load balance, because the full load is twice the present load. The normal equations $(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}} r$ give the Gauss–Newton step. The table below lists the iterations from the starting state (23°C, 28°C, $\ln 6$), the forward solution of both balances at $G = 6\,\text{kW/K}$: the columns are the iteration, the three entries and the objective.

| iteration [–] | $T_1$ [°C] | $T_2$ [°C] | $g$ [–] | $\mathcal{L}$ [–] |
|---|---|---|---|---|
| 0 | 23.0000 | 28.0000 | 1.7918 | 1.281 |
| 1 | 22.5998 | 27.1996 | 1.8718 | 6.083 |
| 2 | 22.6154 | 27.2308 | 1.8718 | $6 \times 10^{-18}$ |
| 3 | 22.6154 | 27.2308 | 1.8718 | $2 \times 10^{-26}$ |

At the start both balances hold, so the objective 1.281 is the prior row's alone, $\frac{1}{2}((\ln 6 - \ln 6.5)/0.05)^2$. The row on $g$ is linear, so the first step puts $g$ at $\hat{g}$; the objective rises at that step because the step is taken on the linearization of $\mathrm{e}^{-g}$, and the balance residuals it leaves, $-0.016$ and $-0.031$ K, are magnified a hundredfold by the division by $\sigma_p$. The second step closes both balances. At the most probable state

$$\hat{\mathbf{z}} = \begin{bmatrix} 22.615\,^{\circ}\mathrm{C} \\ 27.231\,^{\circ}\mathrm{C} \\ 1.871802 \end{bmatrix}$$

every row is zero, $T_1 = 18 + 30/6.5$ and $T_2 = 18 + 60/6.5$.

For rows divided by their standard deviations, as here, the information matrix is $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$. When the rows are left in their physical units, with Jacobian $\mathbf{J}_u$ and weights $\mathbf{W} = \mathrm{diag}(1/\sigma_k^2)$, it is $\mathbf{J}_u^{\mathsf{T}}\mathbf{W}\mathbf{J}_u$, the same matrix, since $\mathbf{J} = \mathbf{W}^{1/2}\mathbf{J}_u$. Its inverse is the Laplace covariance of the state. The curvature term $\sum_k r_k\, \partial^2 r_k/\partial \mathbf{z}^2$ that Newton's method would add vanishes here, because every row is zero at $\hat{\mathbf{z}}$, so

$$\mathbf{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1} = \begin{bmatrix} 0.05335 & 0.10651 & -0.01154 \\ 0.10651 & 0.21312 & -0.02308 \\ -0.01154 & -0.02308 & 0.00250 \end{bmatrix} \qquad (\sigma_p = 0.01\,\mathrm{K}). \tag{2.32}$$

**The ridge of the posterior.** The eigenvalues of this $\mathbf{\Sigma}$ are 0.2689, $1.0 \times 10^{-4}$ and $9.3 \times 10^{-7}$: one direction carries almost all of the uncertainty, and the other two shrink with $\sigma_p^2$. In the limit $\sigma_p \to 0$ the covariance is a single outer product of the vector $\mathbf{v} = (-Q_1/G,\ -Q_2/G,\ 1)^{\mathsf{T}}$,

$$\mathbf{\Sigma} \to \sigma_g^2\, \mathbf{v}\mathbf{v}^{\mathsf{T}} = \begin{bmatrix} 0.05325 & 0.10651 & -0.01154 \\ 0.10651 & 0.21302 & -0.02308 \\ -0.01154 & -0.02308 & 0.00250 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} -4.615 \\ -9.231 \\ 1 \end{bmatrix}, \tag{2.33}$$

where $\mathbf{v}$ is the direction along which both balance rows stay zero (it is orthogonal to the first two rows of $\mathbf{J}$) and $\sigma_g^2$ the variance the third row leaves along it. The three entries move as one along $\mathbf{v}$: a unit rise of $g$ cools the hall by 4.615 K at the present load and 9.231 K at full load. The limit is reached quickly: $\sqrt{\Sigma_{22}}$ is 1.1014 K at $\sigma_p = 1$ K, 0.4722 K at 0.1 K, 0.46165 K at 0.01 K, 0.46154 K at 0.001 K and 0.461538 K at $10^{-4}$ K, and $\sqrt{\Sigma_{33}} = 0.05$ at every $\sigma_p$. The answers below are read off the limit.

The fit supplies the objects that question A was reduced to. The log-conductance is the third entry of the state, $\hat{z}_3 = 1.871802$ with $\sqrt{\Sigma_{33}} = 0.05$, and the conductance has the gradient and band

$$\mathbf{g} = \frac{\partial G}{\partial \mathbf{z}} = \begin{bmatrix} 0 \\ 0 \\ G \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 6.5 \end{bmatrix} \mathrm{kW/K}, \qquad \sqrt{\mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}} = 6.5 \times 0.05 = 0.325\,\mathrm{kW/K},$$

so that the reported answer is $G = 6.5 \pm 0.325$ kW/K, with the central 95 % band $6.5 \pm 1.960 \times 0.325$, from 5.863 to 7.137 kW/K. The exact lognormal summaries follow from the normal of the third entry and stand beside it: the median is $G_{\mathrm{ref}}\, \mathrm{e}^{\hat{z}_3} = 6.5$ kW/K, the location above; the mean, by the moment-generating function of the normal, is $G_{\mathrm{ref}}\, \mathrm{e}^{\hat{z}_3 + \Sigma_{33}/2} = 6.5\, \mathrm{e}^{0.00125} = 6.508$ kW/K;

the standard deviation is 0.3256 kW/K, a gap of 0.0006 from the band; and the central 95 % credible interval is $6.5\,\mathrm{e}^{\pm 1.960\times 0.05}$, from 5.893 to 7.169 kW/K, whose ends lie 0.030 and 0.032 kW/K above those of the band. Each of these is a summary of a distribution; the answer to question A is the distribution itself, equation (2.9) with $f(\mathbf{z}) = G$, and any single number reported from it must say which summary it is. The process gain $K_Q = 1/G$ is a decreasing function of the same entry, so its quantiles are the images of these in reverse order.

### 2.4.2 The threshold probability of the hall at full load

Question B asks whether the hall will remain below $T_{\max} = 27.5°\mathrm{C}$ at full load, that is, for

$$\Pr(T_2 < 27.5°\mathrm{C} \mid \mathbf{y}).$$

It is a prediction under a threshold, of the general form (2.12). The requirement is the upper limit $T_2 < T_{\max}$ on the second entry of the state, so the form applies with $f(\mathbf{z}) = -T_2 = -z_2$ and $f^* = -T_{\max}$; equivalently, the probability of the requirement is one minus that of the exceedance, $1 - \Pr(T_2 \geq T_{\max} \mid \mathbf{y})$, and the assurance rule is applied to the requirement. The plant enters through the scenario node: the event is a bound on an entry that no sensor reads, tied to the third entry by the full-load balance row. The tail is read from the fit. Under the Laplace approximation the second entry is normal with center $\hat{z}_2$ and variance $\Sigma_{22}$, the second diagonal entry of $\boldsymbol{\Sigma}$, so the probability is a normal tail in those two numbers, read off $\hat{\mathbf{z}}$ and the limit of equation (2.33), and this is the value reported:

$$\Pr(T_2 < T_{\max} \mid \mathbf{y}) \approx \Phi\!\left(\frac{T_{\max} - \hat{z}_2}{\sqrt{\Sigma_{22}}}\right) = \Phi\!\left(\frac{27.5 - 27.231}{0.4615}\right) = \Phi(0.5833) = 0.720. \tag{2.34}$$

**The exact value along the ridge.** An exact value stands beside it, through the ridge. Along $\mathbf{v}$ the full-load temperature is a decreasing function of the third entry, $T_2 = T_c + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$, so the bound on $T_2$ is a bound on $g$, at the conductance where the second row meets the limit,

$$\Pr(T_2 < 27.5°\mathrm{C} \mid \mathbf{y}) = \Pr(G > G^* \mid \mathbf{y}), \qquad G^* = \frac{Q_2}{T_{\max} - T_c} = 6.3158\,\mathrm{kW/K}, \tag{2.35}$$

that is $g > \ln(G^*/G_{\mathrm{ref}}) = 1.843053$; and the third entry is exactly normal, so

$$\Pr(T_2 < 27.5°\mathrm{C} \mid \mathbf{y}) = \Phi\!\left(\frac{\hat{z}_3 - \ln(G^*/G_{\mathrm{ref}})}{\sqrt{\Sigma_{33}}}\right) = \Phi\!\left(\frac{1.871802 - 1.843053}{0.05}\right), \tag{2.36}$$

$$\Pr(T_2 < 27.5°\mathrm{C} \mid \mathbf{y}) = \Phi(0.5750) = 0.717. \tag{2.37}$$

The normal tail of the entry and the exact value differ by 0.003, because $T_2$ is not linear in $g$ along the ridge. The probability that the hall stays below the limit at full load is 72.0 %, with 71.7 % exact beside it, and the probability that it exceeds the limit is 28.0 %, with 28.3 % exact, conditional on the row of equation (2.29), the operating conditions and the steady-state thermal model. Whether a 28 % probability of exceeding the limit is acceptable is not a probability question; it depends on the consequences of exceedance and the operating criteria that apply. The assurance rule of equation (2.12) states such a criterion as a number, applied to the probability that the requirement holds: at an assurance level $\kappa = 0.95$ the requirement $T_2 < 27.5°\mathrm{C}$ is not accepted by either value, since 0.720 and 0.717 both lie below 0.95.

Figure 2.2 shows the two questions side by side. In panel (a) the horizontal axis is the conductance $G$ and the vertical axis its posterior density; the shaded area to the right of the dashed line at $G^*$ is the exact 0.717. In panel (b) the horizontal axis is the full-load temperature $T_2$ and the vertical axis its posterior density; the solid curve is the exact density implied by the lognormal $G$, the dashed curve the normal of the entry of equation (2.34), whose area to the

left of the limit is the reported 0.720, marked on the panel, and the shaded area under the solid curve to the left of the limit is the same exact 0.717. The two shaded areas are one probability, seen once on the conductance and once on the temperature: the event $T_2 < 27.5°\mathrm{C}$ and the event $G > G^*$ are the same set of states.

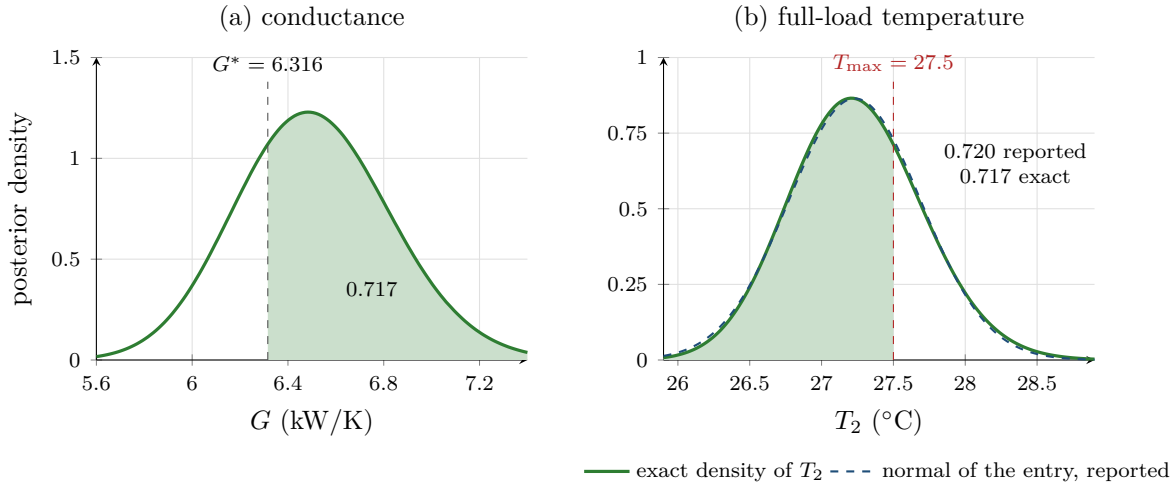



Figure 2.2: One posterior, two questions. (a) Lognormal posterior of the conductance, density $(G\,s\sqrt{2\pi})^{-1}\exp[-(\ln G - \ln 6.5)^2/(2s^2)]$ with $s = \sqrt{\Sigma_{33}} = 0.05$; shaded, $\Pr(G > 6.316\,\mathrm{kW/K} \mid \mathbf{y}) = 0.717$. (b) Posterior of the full-load temperature $T_2 = 18 + 60/G$, lognormal in $T_2 - 18$ with median 9.231 K and log-spread 0.05 (solid), and the normal of the entry, $\mathcal{N}(\hat{z}_2, \Sigma_{22}) = \mathcal{N}(27.231, 0.4615^2)$ (dashed), whose tail is the reported 0.720; shaded, the exact $\Pr(T_2 < 27.5°\mathrm{C} \mid \mathbf{y}) = 0.717$.

**The check against the exact value and against draws.** Equation (2.37) is exact under the normal posterior of $g$, because the map from $g$ to $T_2$ is monotone along the ridge; equation (2.34) is the route that every derived quantity has. Had $T_2$ not been carried in the state, the same band would come from the gradient of $f(\mathbf{z}) = T_c + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$, $\mathbf{g} = (0,\ 0,\ -9.231)^{\mathsf{T}}$ K, and it agrees with $\Sigma_{22}$ because in the limit of equation (2.33) $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} = \sigma_g^2(\mathbf{g}^{\mathsf{T}}\mathbf{v})^2$ and $\mathbf{g}$ has the same projection on $\mathbf{v}$, $-9.231$, as the unit vector on $T_2$. Table 2.6 compares the reported and the exact probabilities for the limit of the example and three others. The first column is the limit, the second the equivalent conductance threshold of equation (2.35), the third the standardized distance in $g$, and the last two the reported normal of the entry, with mean $\hat{z}_2$ and variance $\Sigma_{22}$, and the exact probability. The two differ by at most 0.007 over these limits and by 0.003 at 27.5°C, because the curvature of $60/G$ across a posterior of 5 % spread is small; panel (b) of Figure 2.2 shows the slight right skew the normal misses. Small as it is, the gap can cross a decision level: at $T_{\max} = 28°\mathrm{C}$ and $\kappa = 0.95$ the normal gives 0.9522, which accepts the requirement, and the exact value 0.9453, which does not. Both values are then stated, and the engineer sees that the requirement sits on the level. A second check draws 200 000 values of $g$ from its posterior, computes $T_2$ for each, and counts the fraction below the limit: 0.7175, against the exact 0.7174, within the sampling error of 0.001.

**What the two questions share.** Both questions used the same posterior. Estimation, $p(G \mid \mathbf{y})$, describes the uncertainty about the conductance; the threshold probability, $\Pr(T_2 < T_{\max} \mid \mathbf{y})$, integrates that uncertainty over the states that satisfy the temperature limit. Neither required a new posterior; each asked a different question of the same one. That reuse is one of the principal advantages of a probabilistic model of the plant: the expensive step, the posterior, is paid once, and each new question is an integral over it.

Table 2.6: The threshold probability $\Pr(T_2 < T_{\max} \mid \mathbf{y})$ at $Q_2 = 60\,\mathrm{kW}$, $T_c = 18^\circ\mathrm{C}$, $g \mid \mathbf{y} \sim \mathcal{N}(\ln 6.5, 0.05^2)$: reported, by the normal $\mathcal{N}(\hat{z}_2, \Sigma_{22})$ of the entry $T_2$ (equation (2.34)), and exact (equation (2.36)).

| $T_{\max}$ [°C] | $G^* = Q_2/(T_{\max} - T_c)$ [kW/K] | $(\hat{z}_3 - \ln G^*)/\sqrt{\Sigma_{33}}$ [–] | reported, normal of $T_2$ [–] | exact [–] |
|---|---|---|---|---|
| 27.0 | 6.6667 | −0.5064 | 0.3085 | 0.3063 |
| 27.5 | 6.3158 | 0.5750 | 0.7202 | 0.7174 |
| 28.0 | 6.0000 | 1.6009 | 0.9522 | 0.9453 |
| 28.5 | 5.7143 | 2.5767 | 0.9970 | 0.9950 |

### 2.4.3 The posterior of the loop's trip rate

A different record of the same hall: the log of its discharge loop's trips, each one valve stiction sending the loop to manual. Over the last 1000 h of logged operation, about six weeks, the loop tripped twice, and the engineer asks for the probability that the coming twelve-hour night run passes without a trip. This is a prediction question, not only a question about the trip rate. Let $\lambda$ be the unknown trip rate, in trips per hour, $k = 2$ the observed number of trips, $T_{\mathrm{tot}} = 1000\,\mathrm{h}$ the logged operation behind the count, and $\tau = 12\,\mathrm{h}$ the length of the run. The trips are taken to follow a homogeneous Poisson process, so that the count in the logged time is Poisson with mean $\lambda T_{\mathrm{tot}}$,

$$N(T_{\mathrm{tot}}) \mid \lambda \sim \mathrm{Poisson}(\lambda T_{\mathrm{tot}}), \qquad \Pr(k \mid \lambda, T_{\mathrm{tot}}) = \frac{(\lambda T_{\mathrm{tot}})^k \, \mathrm{e}^{-\lambda T_{\mathrm{tot}}}}{k!}.$$

This is a modeling choice. It is what lets the log be summarized by $k$ and $T_{\mathrm{tot}}$ alone, and it fails when one stuck valve produces several trips (a loop returned to automatic and sent straight back to manual) or when the trips crowd into the hours of large valve moves; the count here is of trips after the supervisor's re-arm delay, which merges repeats into one. As a function of $\lambda$, the likelihood is proportional to $\lambda^k \, \mathrm{e}^{-\lambda T_{\mathrm{tot}}}$; the factor $T_{\mathrm{tot}}^k/k!$ does not depend on the rate. The prior is a gamma distribution in the shape–rate parameterization, $\lambda \sim \mathrm{Gamma}(a, b)$, with density proportional to $\lambda^{a-1} \, \mathrm{e}^{-b\lambda}$ and, here,

$$a = 2, \qquad b = 500\,\mathrm{h},$$

the trip records of the site's sister loops, whose prior mean $a/b = 0.004\,\mathrm{h}^{-1}$ is one trip every 250 h. Bayes' rule gives the posterior up to a constant as likelihood times prior, $p(\lambda \mid k, T_{\mathrm{tot}}) \propto \Pr(k \mid \lambda, T_{\mathrm{tot}}) \, p(\lambda)$.

The rate enters the state, as every rate does in this book, through its log-rate $\omega = \ln(\lambda/\lambda_{\mathrm{ref}})$ with $\lambda_{\mathrm{ref}} = 0.001\,\mathrm{h}^{-1}$, so that the state is the one number

$$\omega = \ln \frac{\lambda}{\lambda_{\mathrm{ref}}}, \qquad \lambda = \lambda_{\mathrm{ref}} \, \mathrm{e}^{\omega}.$$

In the general forms $\mathbf{z}$ is here the one number $\omega$, and every quantity of the fit is a number. The rate is asked for first, an estimation question of the general form (2.9) with $f(\omega) = \lambda = \lambda_{\mathrm{ref}} \, \mathrm{e}^{\omega}$, increasing in $\omega$. It is read from the MAP log-rate $\hat{\omega}$, its variance $\sigma_\omega^2$ and the derivative $\mathrm{d}\lambda/\mathrm{d}\omega = \hat{\lambda}$, whose band is $\hat{\lambda} \, \sigma_\omega$; the exact gamma posterior that the objective identifies is set beside it. Neither row is Gaussian, and each enters the objective as its negative log-probability term. The count row is the negative log of the Poisson probability; the prior row is the negative log of the gamma density carried over to $\omega$, where the change of variable multiplies the density by

$\mathrm{d}\lambda/\mathrm{d}\omega = \lambda$ and so adds $-\ln\lambda$ (the rule, that a density carried to a new coordinate is multiplied by the derivative of the old coordinate with respect to the new, is derived in Section 4.3.3):

$$\mathcal{L}(\omega) = \underbrace{T_{\mathrm{tot}}\lambda - k\ln\lambda}_{\text{count row, } \mathcal{L}_c} + \underbrace{b\lambda - a\ln\lambda}_{\text{prior row, } \mathcal{L}_0} + \text{const.} \tag{2.38}$$

With primes for derivatives in $\omega$ and $\mathrm{d}\lambda/\mathrm{d}\omega = \lambda$, each row has a slope and a curvature, and they add:

$$\mathcal{L}_c' = T_{\mathrm{tot}}\lambda - k, \quad \mathcal{L}_c'' = T_{\mathrm{tot}}\lambda, \qquad \mathcal{L}_0' = b\lambda - a, \quad \mathcal{L}_0'' = b\lambda, \qquad \mathcal{L}' = \mathcal{L}_c' + \mathcal{L}_0', \quad \mathcal{L}'' = \mathcal{L}_c'' + \mathcal{L}_0''.$$

The Newton step $\Delta\omega = -\mathcal{L}'/\mathcal{L}''$ moves to the minimum of the parabola with the current slope and curvature. The table below lists the iterations from the maximum likelihood rate $k/T_{\mathrm{tot}} = 0.002\,\mathrm{h}^{-1}$, that is $\omega = \ln 2$, where the count row's slope is zero: the columns are the iteration, the log-rate, the rate, the slopes of the two rows, the slope and curvature of the objective, and the step.

| iteration $j$ [–] | $\omega_j$ [–] | $\lambda$ [$\mathrm{h}^{-1}$] | $\mathcal{L}_c'$ [–] | $\mathcal{L}_0'$ [–] | $\mathcal{L}'$ [–] | $\mathcal{L}''$ [–] | $\Delta\omega$ [–] |
|---|---|---|---|---|---|---|---|
| 0 | 0.69315 | 0.002000 | 0.0000 | −1.0000 | −1.0000 | 3.0000 | 0.33333 |
| 1 | 1.02648 | 0.002791 | 0.7912 | −0.6044 | 0.1868 | 4.1868 | −0.04462 |
| 2 | 0.98186 | 0.002669 | 0.6694 | −0.6653 | 0.0041 | 4.0041 | −0.00103 |
| 3 | 0.98083 | 0.002667 | 0.6667 | −0.6667 | $2\times10^{-6}$ | 4.0000 | $-5\times10^{-7}$ |

The slope vanishes at $\lambda = (a+k)/(b+T_{\mathrm{tot}})$, so the mode is $\hat\omega = \ln(8/3) = 0.98083$ and $\lambda_{\mathrm{ref}}\,\mathrm{e}^{\hat\omega} = 4/1500\,\mathrm{h} = 0.002667\,\mathrm{h}^{-1}$. The curvature there is the information, $\Lambda = \mathcal{L}''(\hat\omega) = T_{\mathrm{tot}}\hat\lambda + b\hat\lambda = 2.667 + 1.333 = 4 = a + k$: the count row contributes the expected number of trips in the logged time, the prior row $b\hat\lambda$, and the two add as the information of independent rows does. Then $\sigma_\omega^2 = 1/\Lambda = 0.25$, the log-rate has $\sigma_\omega = 0.5$, and the rate, a derived quantity with derivative $\mathrm{d}\lambda/\mathrm{d}\omega = \hat\lambda$, has $\hat\lambda\,\sigma_\omega = 0.002667 \times 0.5 = 0.001333\,\mathrm{h}^{-1}$. The reported answer is the trip rate $0.002667 \pm 0.001333\,\mathrm{h}^{-1}$.

The objective also identifies the exact posterior. Back in $\lambda$, without the $-\ln\lambda$ of the change of variable, the negative log posterior is $(b+T_{\mathrm{tot}})\lambda - (a+k-1)\ln\lambda$, the negative log of the kernel $\lambda^{a+k-1}\,\mathrm{e}^{-(b+T_{\mathrm{tot}})\lambda}$ of a gamma density. Therefore

$$\lambda \mid k, T_{\mathrm{tot}} \sim \mathrm{Gamma}(a', b'), \qquad a' = a + k, \quad b' = b + T_{\mathrm{tot}}. \tag{2.39}$$

For the loop's trips, $a' = 2 + 2 = 4$ and $b' = 500 + 1000 = 1500\,\mathrm{h}$, so $\lambda \mid \mathbf{y} \sim \mathrm{Gamma}(4, 1500\,\mathrm{h})$. The posterior mean and standard deviation are

$$\mathbb{E}[\lambda \mid \mathbf{y}] = \frac{a'}{b'} = \frac{4}{1500\,\mathrm{h}} = 0.002667\,\mathrm{h}^{-1}, \qquad \sqrt{\mathrm{Var}[\lambda \mid \mathbf{y}]} = \frac{\sqrt{a'}}{b'} = 0.001333\,\mathrm{h}^{-1},$$

Set like with like, the reported posterior and the exact one differ, because the Laplace normal of the log-rate makes the rate lognormal: its median is $0.002667\,\mathrm{h}^{-1}$ against the gamma's $0.002448\,\mathrm{h}^{-1}$, a gap of $0.000219\,\mathrm{h}^{-1}$; its mean $\hat\lambda\,\mathrm{e}^{\sigma_\omega^2/2} = 0.003022\,\mathrm{h}^{-1}$ against the gamma's $0.002667\,\mathrm{h}^{-1}$, a gap of $0.000355\,\mathrm{h}^{-1}$; and its central 95 % interval runs from 0.001001 to $0.007105\,\mathrm{h}^{-1}$ against the gamma's 0.000727 to $0.005845\,\mathrm{h}^{-1}$. The pair the Newton fit returned coincides with the gamma's mean and standard deviation: the mode of the log-rate maps onto the gamma mean and the linearized band onto the gamma standard deviation, for every gamma posterior, but that compares different summaries, a median of one posterior with a mean of the other. The gamma mean is one trip every 375 h. It lies between the prior mean $a/b = 0.004\,\mathrm{h}^{-1}$, one every 250 h, and the maximum likelihood estimate $k/T_{\mathrm{tot}} = 0.002\,\mathrm{h}^{-1}$, one every 500 h, as the weighted form $a'/b' = \frac{b}{b'}\,\frac{a}{b} + \frac{T_{\mathrm{tot}}}{b'}\,\frac{k}{T_{\mathrm{tot}}} = \frac{1}{3}(0.004) + \frac{2}{3}(0.002)$ shows. An estimation question has been answered. The engineer's question, a clean night run, has not.

### 2.4.4 The probability of a clean night run

Let $T$ be the time to the next trip of the loop, counted from the start of the run. (In this example $T$, in hours, is a time; in the hall example, $T_1$ and $T_2$ are hall temperatures in °C. The units distinguish them.) Under the Poisson process the time to the next trip does not depend on how long ago the last one was, so the question is

$$\Pr(T > \tau \mid \mathbf{y}), \qquad \tau = 12\,\mathrm{h}. \tag{2.40}$$

For a known constant rate the probability of no trip is exponential, $\Pr(T > \tau \mid \lambda) = \mathrm{e}^{-\lambda\tau}$. The rate is not known, and the event is a future one that stays random even when the rate is known, so the question is a prediction of the general form (2.13), in its event form (2.8): $y_{\mathrm{new}}$ is the event $T > \tau$, and $p(y_{\mathrm{new}} \mid \mathbf{z}) = \mathrm{e}^{-\lambda\tau}$ with $\lambda = \lambda_{\mathrm{ref}}\,\mathrm{e}^{\omega}$. The integral is a function of the one number $\omega$ averaged over its posterior, and under the Laplace approximation that posterior is $\mathcal{N}(\hat{\omega}, \sigma_\omega^2)$, so the value read from the state is a normal expectation in $\hat{\omega}$ and $\sigma_\omega^2$:

$$\Pr(T > \tau \mid \mathbf{y}) \approx \mathbb{E}_{\omega\sim\mathcal{N}(\hat{\omega},\sigma_\omega^2)}\big[\mathrm{e}^{-\lambda_{\mathrm{ref}}\mathrm{e}^{\omega}\tau}\big] \approx \sum_{j=1}^{5} w_j\,\mathrm{e}^{-\lambda_{\mathrm{ref}}\mathrm{e}^{\omega_j}\tau}, \qquad \omega_j = \hat{\omega} + \sqrt{2}\,\sigma_\omega\,x_j. \tag{2.41}$$

The sum is the five-point Gauss–Hermite rule, which replaces a normal expectation by a weighted sum at fixed nodes $x_j$ with weights $w_j$ and is exact for a polynomial of degree nine. The table below evaluates it for the night run, with $\hat{\omega} = 0.98083$ and $\sigma_\omega = 0.5$: the columns are the node, its weight, the log-rate and rate there, the probability of no trip at that rate, and the weighted term.

| $x_j$ [–] | $w_j$ [–] | $\omega_j$ [–] | $\lambda_j$ [$\mathrm{h}^{-1}$] | $\mathrm{e}^{-12\lambda_j}$ [–] | term [–] |
|---|---|---|---|---|---|
| −2.02018 | 0.01126 | −0.44766 | 0.000639 | 0.99236 | 0.01117 |
| −0.95857 | 0.22208 | 0.30302 | 0.001354 | 0.98388 | 0.21850 |
| 0 | 0.53333 | 0.98083 | 0.002667 | 0.96851 | 0.51654 |
| 0.95857 | 0.22208 | 1.65864 | 0.005252 | 0.93892 | 0.20851 |
| 2.02018 | 0.01126 | 2.40931 | 0.011126 | 0.87501 | 0.00985 |

The terms sum to 0.9646, and a direct numerical integration agrees to six decimals. The predicted probability of a night run without a trip is 96.46 %, and the probability of at least one trip is 0.0354; these are the values reported.

The exact value stands beside them, because the objective identified the posterior as a gamma and the same integral then has a closed form. The conditional probability of a clean run is averaged over the posterior of the rate,

$$\Pr(T > \tau \mid \mathbf{y}) = \int_0^\infty \mathrm{e}^{-\lambda\tau}\,p(\lambda \mid \mathbf{y})\,\mathrm{d}\lambda.$$

Substitute the density of the gamma posterior of equation (2.39),

$$p(\lambda \mid \mathbf{y}) = \frac{b'^{\,a'}}{\Gamma(a')}\,\lambda^{a'-1}\,\mathrm{e}^{-b'\lambda},$$

and combine the two exponentials:

$$\Pr(T > \tau \mid \mathbf{y}) = \frac{b'^{\,a'}}{\Gamma(a')}\int_0^\infty \lambda^{a'-1}\,\mathrm{e}^{-(b'+\tau)\lambda}\,\mathrm{d}\lambda.$$

The remaining integral is the standard gamma integral, $\int_0^\infty \lambda^{a'-1}\mathrm{e}^{-c\lambda}\,\mathrm{d}\lambda = \Gamma(a')/c^{a'}$, obtained from the definition of the gamma function by the substitution $x = c\lambda$; with $c = b' + \tau$ the gamma functions cancel, and

$$\Pr(T > \tau \mid \mathbf{y}) = \left(\frac{b'}{b'+\tau}\right)^{a'}. \tag{2.42}$$

For the night run,

$$\Pr(T > 12\,\mathrm{h} \mid \mathbf{y}) = \left(\frac{1500}{1512}\right)^4 = 0.9686,$$

and the probability of at least one trip is 0.0314. The calculation needed no Gaussian or Laplace approximation: with a gamma posterior and an exponential probability of no trip, the predictive probability has an exact closed form. The gap between the two is 0.00406 for the run, and over the next 100 h, about eight night runs, the state gives 0.748 against the exact 0.7725, a gap of 0.024. The state route errs low because the lognormal it implies for $\lambda$ has the larger mean $0.00302\,\mathrm{h}^{-1}$ and a heavier upper tail than the gamma. No decision of this chapter sits between the two values; an assurance level of 0.965 on the run would, and both would then be stated, the one read from the state failing it and the exact one meeting it.

### 2.4.5 The error of the plug-in trip rate

A common shortcut substitutes the posterior mean rate into the exponential:

$$\Pr_{\text{plug}}(T > 12\,\mathrm{h}) = \mathrm{e}^{-\mathbb{E}[\lambda|\mathbf{y}]\times 12\,\mathrm{h}} = \mathrm{e}^{-0.0320} = 0.9685,$$

against the exact predictive 0.9686; over the next 100 h it gives $\mathrm{e}^{-0.2667} = 0.7659$ against 0.7725. The two differ because averaging and exponentiation do not commute,

$$\mathbb{E}\big[\mathrm{e}^{-\lambda\tau} \mid \mathbf{y}\big] \neq \mathrm{e}^{-\mathbb{E}[\lambda|\mathbf{y}]\,\tau} \quad \text{in general,} \tag{2.43}$$

since the probability of no trip is a nonlinear function of the rate. The direction of the difference follows from the shape of that function. For $\tau > 0$ the map $\lambda \mapsto \mathrm{e}^{-\lambda\tau}$ has second derivative $\tau^2\mathrm{e}^{-\lambda\tau} > 0$, so it is convex, and Jensen's inequality, the standard result that the expectation of a convex function is at least the function of the expectation, gives $\mathbb{E}[\mathrm{e}^{-\lambda\tau} \mid \mathbf{y}] \geq \mathrm{e}^{-\mathbb{E}[\lambda|\mathbf{y}]\tau}$: the plug-in value understates the exact probability of a clean run. The low rates the posterior admits raise that probability by more than the high rates lower it. In the log-rate the same function is concave while $\lambda\tau < 1$, so the reported average (2.41) starts below the plug-in and crosses it at 356 h. For one run the plug-in sits 0.0001 from the exact value, closer than the reported 0.9646. The plug-in value is wrong not because it is far from the predictive value, but because it answers a different question: the probability of a clean run for a loop whose trip rate is known to be $0.002667\,\mathrm{h}^{-1}$, which is not the loop in question, and it carries no band.

Table 2.7 shows how the values change with the horizon. The first column is the horizon; the second the reported average of equation (2.41); the third the exact predictive probability of no trip, equation (2.42); the fourth their gap; the fifth the plug-in value at the posterior mean; the sixth the exact value minus the plug-in; and the last the plug-in value at the maximum likelihood estimate $0.002\,\mathrm{h}^{-1}$, which ignores the prior as well as the uncertainty. For one run or one day the exact and plug-in values agree to three decimals. Their ratio grows with the horizon, from 1.0005 at 24 h to 1.87 at 1000 h and 7.0 at 2000 h; the absolute difference peaks near 800 h, at 0.0625, and then falls, to 0.0289 at 2000 h, because both values tend to zero; the gap of the reported average peaks in the hundreds of hours, at 0.0387 among the rows of the table. Figure 2.3 draws the ingredients. In panel (a) the horizontal axis is the trip rate in trips per 1000 h and the vertical axis the probability density per unit of that axis; the blue curve is the prior, the green curve the posterior, and the dotted lines mark the maximum likelihood estimate, the posterior mean and the prior mean. In panel (b) the horizontal axis is the horizon and the vertical axis the probability of no trip; the green curve is the exact predictive, the blue dash-dotted curve the reported average over the normal of the log-rate and the dashed red curve the plug-in, and the vertical dotted line marks the 12 h run. The exact curve lies above the plug-in curve at every horizon, and their ratio grows as the horizon lengthens, because the longer the horizon, the more the probability of no trip bends across the spread of the posterior; the blue curve falls below both at first and crosses the plug-in at 356 h.

Table 2.7: Probability of no trip of the discharge loop over a horizon $\tau$: the reported average over the normal of the log-rate (equation (2.41), $\hat{\omega} = \ln(8/3)$, $\sigma_\omega = 0.5$); exact, from the posterior Gamma$(4, 1500\,\text{h})$, $(1500/(1500+\tau))^4$; plug-in $e^{-4\tau/1500}$; plug-in at the maximum likelihood estimate $e^{-0.002\tau}$. Gaps are computed before rounding.

| $\tau$ [h] | reported, from the state [–] | exact [–] | gap [–] | plug-in, posterior mean [–] | exact minus plug-in [–] | plug-in, maximum likelihood [–] |
|---|---|---|---|---|---|---|
| 12 | 0.9646 | 0.9686 | 0.0041 | 0.9685 | 0.0001 | 0.9763 |
| 24 | 0.9307 | 0.9385 | 0.0078 | 0.9380 | 0.0005 | 0.9531 |
| 100 | 0.7481 | 0.7725 | 0.0244 | 0.7659 | 0.0065 | 0.8187 |
| 500 | 0.2777 | 0.3164 | 0.0387 | 0.2636 | 0.0528 | 0.3679 |
| 1000 | 0.1015 | 0.1296 | 0.0281 | 0.0695 | 0.0601 | 0.1353 |
| 2000 | 0.0205 | 0.0337 | 0.0132 | 0.0048 | 0.0289 | 0.0183 |

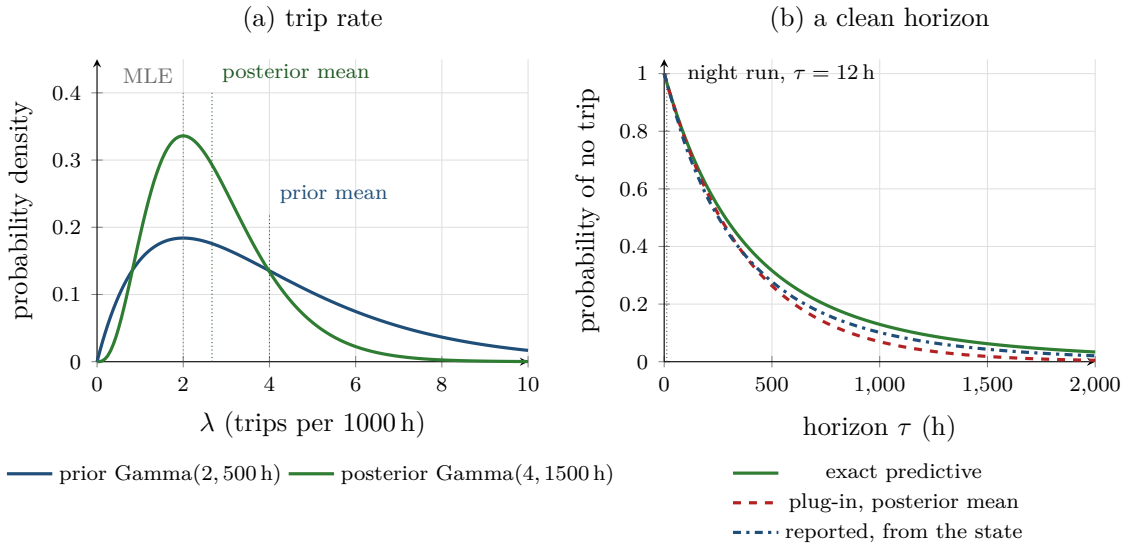



Figure 2.3: The loop's trip rate and the probability of a clean horizon. (a) Prior $0.25\,x\,e^{-0.5x}$ and posterior $(1.5^4/3!)\,x^3e^{-1.5x}$ in $x = 1000\lambda$; dotted lines at the maximum likelihood estimate 2.0, the posterior mean 2.67 and the prior mean 4.0 trips per 1000 h. (b) Exact predictive probability of no trip $(1500/(1500+\tau))^4$ of equation (2.42), plug-in $e^{-4\tau/1500}$ and the reported average of equation (2.41); at the 12 h run they are 0.9686, 0.9685 and 0.9646, at 100 h 0.7725, 0.7659 and 0.7481.

**Principle 2.2.** *Predicting a clean run requires averaging the probability of no trip over the uncertainty in the trip rate. Substituting a point estimate of the rate answers the question for a different loop, one whose rate is known.*

### 2.4.6 Two loops clean through the night as a joint event

A site has two other discharge loops, A and B, each with its own valve, and one engineer on call for both. Figure 2.4 draws the arrangement: the white box on the left is the one chilled-water pump behind both valves; the two green boxes are the loops, each with its own causes of a trip (stiction in its valve, a load step, a fault in its sensor); the white box on the right is the engineer on call, who receives both trips; the solid arrows are the paths along which chilled water and trips pass. Define the events

$$A_A = \{\text{loop A trips in the night}\}, \qquad A_B = \{\text{loop B trips in the night}\},$$

and write $\bar{A}_A$ and $\bar{A}_B$ for their complements, a clean night on each loop. The engineer wants a clean night on both, the intersection $\bar{A}_A \cap \bar{A}_B$, and fears the opposite corner, both loops tripping in the same night, $A_A \cap A_B$, two responses one engineer cannot give at once. The two are tied by inclusion and exclusion, which holds without any independence assumption:

$$\Pr(\bar{A}_A \cap \bar{A}_B) = 1 - \Pr(A_A) - \Pr(A_B) + \Pr(A_A \cap A_B). \tag{2.44}$$

Each is the general form (2.12) with the indicator of an intersection in place of the indicator of one bound, here with the marginal probabilities given directly rather than computed from a posterior, so there is no state route to report and no gap to state; every probability of this example is conditional on the same evidence, and the bar is left out below to keep the products readable. Suppose $\Pr(A_A) = 0.05$ and $\Pr(A_B) = 0.03$, illustrative values of the size the trip log of Section 2.4.4 gives one loop, 0.0354 for a night and 0.0314 exact. If the two trips are independent, the probability that both occur is the product of the marginal probabilities,

$$\Pr(A_A \cap A_B) = \Pr(A_A)\,\Pr(A_B) = 0.05 \times 0.03 = 0.0015,$$

and the probability of a clean night on both loops is, by equation (2.44) or as the product of the two complements,

$$\Pr(\bar{A}_A \cap \bar{A}_B) = 0.95 \times 0.97 = 0.9215, \tag{2.45}$$

or 92.15 %.

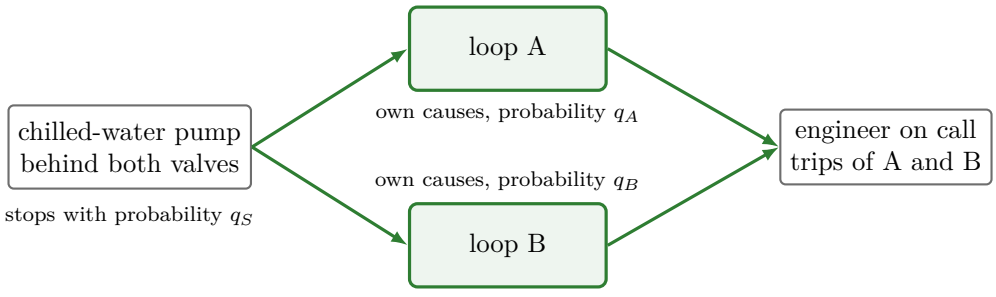



Figure 2.4: Two loops behind one chilled-water pump. Each loop's trip reaches the engineer on call; a clean night on both needs neither to trip. A pump stop (probability $q_S$ in the night) takes the chilled water from both coils and trips both loops at once; the loops' own causes (probabilities $q_A$, $q_B$) are independent given a running pump. With $q_S = 0$ the two trips are independent.

### 2.4.7 One chilled-water pump behind both valves

Both valves draw their chilled water through the same pump. Then independence may not hold, and the probability of the intersection in the form (2.12) must be written with the product rule, which holds without any independence assumption:

$$\Pr(A_A \cap A_B) = \Pr(A_A)\,\Pr(A_B \mid A_A). \tag{2.46}$$

The conditional probability $\Pr(A_B \mid A_A)$ may be much larger than the marginal $\Pr(A_B)$: a trip of loop A raises the probability that the pump has stopped, and a stopped pump takes the chilled water from loop B as well. A shared pump is a common cause that can make the two trips strongly dependent.

An illustrative model shows by how much. Let the pump stop during the night with probability $q_S$, which trips both loops, and let each loop otherwise trip from its own causes, independently, with probabilities $q_A$ and $q_B$. The law of total probability over the two states of the pump, the integral of the form over a state with two values, gives the marginal and the joint probabilities, exactly, with no second route beside them:

$$\Pr(A_A) = q_S + (1 - q_S)\,q_A, \qquad \Pr(A_B) = q_S + (1 - q_S)\,q_B,$$

$$\Pr(A_A \cap A_B) = q_S \cdot 1 + (1 - q_S)\,q_A q_B, \qquad \Pr(\bar{A}_A \cap \bar{A}_B) = (1 - q_S)(1 - q_A)(1 - q_B). \tag{2.47}$$

Holding the marginals at 0.05 and 0.03, the own-cause probabilities are $q_A = (0.05 - q_S)/(1 - q_S)$ and $q_B = (0.03 - q_S)/(1 - q_S)$. For $q_S = 0.01$, $q_A = 0.04/0.99 = 0.040404$ and $q_B = 0.02/0.99 = 0.020202$, so

$$\Pr(A_A \cap A_B) = 0.01 + 0.99 \times 0.040404 \times 0.020202 = 0.01 + 0.000808 = 0.010808,$$

and equation (2.46) gives $\Pr(A_B \mid A_A) = 0.010808/0.05 = 0.216$, seven times the marginal 0.03. The probability of a clean night on both loops is $0.99 \times 0.959596 \times 0.979798 = 0.930808$, the same as $1 - 0.08 + 0.010808$ by equation (2.44). Table 2.8 repeats the calculation for five pump-stop probabilities. The first column is $q_S$; the next two the own-cause probabilities that keep the marginals fixed; then the probability that both loops trip, the conditional probability of the second trip given the first, the ratio of the joint probability to the independent value 0.0015, and the probability of a clean night on both loops. With the marginals unchanged, a stop probability of 0.01 multiplies the probability of a night with both loops tripping by 7.2, and one of 0.02 by 13.5: the pump, not the loops' own conditions, then governs it. The last two rows, a stop in one night of a hundred, 3.65 stops a year, and in one of fifty, are high for a maintained pump and are there to show the trend.

Table 2.8: Two loops behind one chilled-water pump through a night, with the marginals held at $\Pr(A_A) = 0.05$ and $\Pr(A_B) = 0.03$ (illustrative model of equation (2.47)).

| $q_S$ [–] | $q_A$ [–] | $q_B$ [–] | $\Pr(A_A \cap A_B)$ [–] | $\Pr(A_B \mid A_A)$ [–] | ratio to 0.0015 [–] | $\Pr(\bar{A}_A \cap \bar{A}_B)$ [–] |
|---|---|---|---|---|---|---|
| 0 | 0.05000 | 0.03000 | 0.001500 | 0.0300 | 1.00 | 0.92150 |
| 0.001 | 0.04905 | 0.02903 | 0.002422 | 0.0484 | 1.61 | 0.92242 |
| 0.005 | 0.04523 | 0.02513 | 0.006131 | 0.1226 | 4.09 | 0.92613 |
| 0.01 | 0.04040 | 0.02020 | 0.010808 | 0.2162 | 7.21 | 0.93081 |
| 0.02 | 0.03061 | 0.01020 | 0.020306 | 0.4061 | 13.54 | 0.94031 |

The shared pump moves the two corners in opposite directions. The clean night on both loops becomes slightly more probable, 0.9308 against 0.9215, because when the trips come they tend to come together; the night with both loops tripping becomes 7.2 times as probable. A

product of marginals gets the first nearly right and the second badly wrong, and the second is the one that leaves one engineer facing two loops at once. The wording "both loops" names an intersection of events, and an intersection directs the formulation to joint probability and to the modeling of dependence, not to a product of marginals. In a larger plant the event of concern is often a union of many overlapping events, some inlet somewhere above its limit, and a plan over a horizon asks for such a joint tail, which Chapter 18 takes up.

## 2.5 Model comparison, diagnosis and decisions

The questions so far were asked of the state of one model. A question about which model holds or which fault is present is put to several models at once, and its answer is a ratio or a ranking of evidences; a question about which action to take is put to several actions, and its answer is a comparison of expected consequences.

### 2.5.1 Posterior probabilities of competing models

Consider a hall whose temperature readings are inconsistent with its expected behavior, and two hypotheses about it,

$$M_0 : \text{the air path operates normally,} \qquad M_1 : \text{the air path is partially obstructed,}$$

the second a clogged filter or a blocked return, say. The question is which explanation the evidence supports better. This is not a threshold probability on a fixed, known model: the engineer is comparing models. It is a hypothesis comparison, of the general form (2.14) with $M_A = M_1$ and $M_B = M_0$. The plant enters through the two drawings: each is the graph of the air path with its own rows, and $M_1$ carries an entry for the obstruction, a reduction of the conductance of the path, that $M_0$ pins at zero. Each evidence is the integral (2.26) over its own model's state, read under the Laplace approximation from that model's own fit, its MAP state $\hat{\mathbf{z}}_j$ and covariance $\mathbf{\Sigma}_j$. The extra entry of $M_1$ is not rewarded merely for fitting the readings better: the evidence multiplies the height of the fit by the posterior volume $\det(\mathbf{\Sigma}_j)^{1/2}$, and the prior density $p(\hat{\mathbf{z}}_j \mid M_j)$ in that height is low when the prior spreads the entry over a wide range. Treat the model index as an unknown taking the values $M_0, M_1, \ldots$, with prior probabilities $\Pr(M_j)$. Bayes' rule for a discrete unknown, with the law of total probability in the denominator, gives

$$\Pr(M_j \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid M_j)\, \Pr(M_j)}{\sum_k p(\mathbf{y} \mid M_k)\, \Pr(M_k)}, \tag{2.48}$$

where $\Pr(M_j)$ is the prior probability of model $M_j$, $\Pr(M_j \mid \mathbf{y})$ its posterior probability, and $p(\mathbf{y} \mid M_j)$ its evidence, equation (2.26). Dividing equation (2.48) for $M_1$ by the same equation for $M_0$ cancels the denominator and gives the odds form,

$$\frac{\Pr(M_1 \mid \mathbf{y})}{\Pr(M_0 \mid \mathbf{y})} = B_{10}\, \frac{\Pr(M_1)}{\Pr(M_0)}, \qquad B_{10} = \frac{p(\mathbf{y} \mid M_1)}{p(\mathbf{y} \mid M_0)} : \tag{2.49}$$

posterior odds equal the Bayes factor (2.14) times the prior odds.

### 2.5.2 A numerical comparison of two explanations

Suppose the evidences are given, as illustrative values, not read from fits, so there is no Laplace value and exact value to set side by side: $p(\mathbf{y} \mid M_0) = 0.01$ and $p(\mathbf{y} \mid M_1) = 0.04$, and the prior probabilities are $\Pr(M_0) = 0.8$ and $\Pr(M_1) = 0.2$. Equation (2.48) gives

$$\Pr(M_1 \mid \mathbf{y}) = \frac{0.04 \times 0.2}{0.04 \times 0.2 + 0.01 \times 0.8} = \frac{0.008}{0.016} = 0.50.$$

The Bayes factor is

$$B_{10} = \frac{p(\mathbf{y} \mid M_1)}{p(\mathbf{y} \mid M_0)} = \frac{0.04}{0.01} = 4, \qquad \ln B_{10} = \ln 4 = 1.386 \text{ nats},$$

and the odds form checks the result: posterior odds $4 \times (0.2/0.8) = 1$, a posterior probability of $1/(1+1) = 0.50$. The evidence is four times more probable under the obstruction model, yet its posterior probability is only 50 %, because the prior probability assigned to obstruction was low. The Bayes factor expresses the relative support that the evidence provides; it is not by itself a posterior model probability, and turning it into one requires the prior probabilities of the models.

Figure 2.5 makes the distinction visible. In panel (a) the two pairs of bars are the prior and the posterior probabilities of the two models: the evidence moves probability from $M_0$ to $M_1$ until they are equal. In panel (b) the horizontal axis is the prior probability of the obstruction model and the vertical axis its posterior probability; the green curve is $\Pr(M_1 \mid \mathbf{y}) = 4\,\Pr(M_1)/(3\,\Pr(M_1) + 1)$, which equation (2.49) gives for $B_{10} = 4$, and the dashed diagonal is the case $B_{10} = 1$, evidence that does not discriminate. The same Bayes factor of 4 turns a prior of 0.05 into 0.174, a prior of 0.2 into 0.50 (the marked point) and a prior of 0.5 into 0.80: the factor is fixed by the evidence, the conclusion is not.

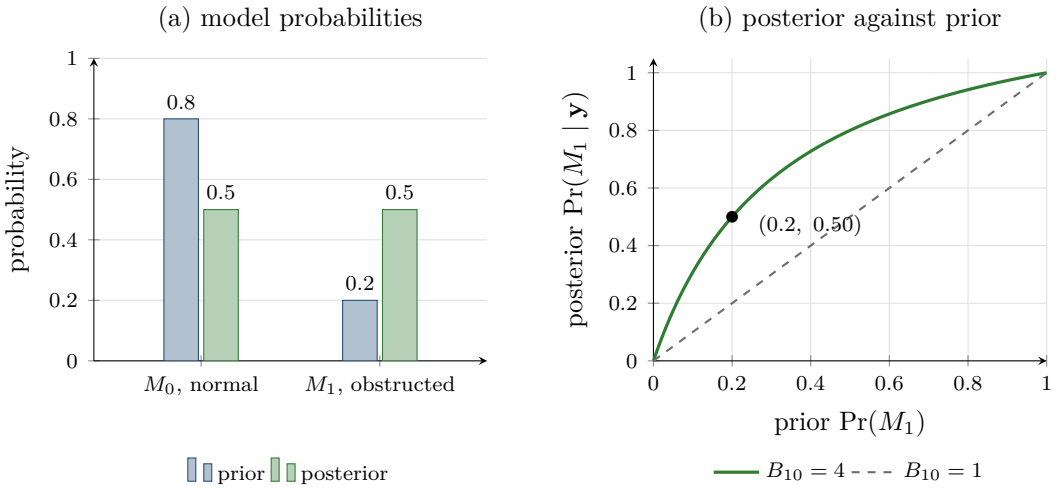



Figure 2.5: Model comparison for the air path of the hall. (a) Prior and posterior probabilities of the normal ($M_0$) and obstructed ($M_1$) models, with evidences 0.01 and 0.04. (b) Posterior probability of $M_1$ as a function of its prior, $4x/(3x+1)$ for a Bayes factor of 4 (equation (2.49)); the dashed diagonal is a Bayes factor of 1; the point is the example.

### 2.5.3 Diagnosis over a family of candidate faults

Fault detection and isolation usually involves many competing explanations, not two. For a hall that runs warmer than its model expects they might be a fouled coil, a fan running slow, a biased hall sensor, a load meter reading low or a floor tile left out of place. The mathematical target is the general form (2.15), normalized as in equation (2.48): the posterior probability of each candidate,

$$\Pr(F_j \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid F_j)\,\Pr(F_j)}{\sum_k p(\mathbf{y} \mid F_k)\,\Pr(F_k)},$$

where $F_j$ is a candidate fault condition, which is the drawing $M_j$ of equation (2.15): the plant graph with that fault written into its rows, so that $p(\mathbf{y} \mid F_j) = p(\mathbf{y} \mid M_j)$. The output is the ranking of the candidates by it, not a single verdict. A fault condition can be represented in two

ways: as a separate model, whose evidence enters equation (2.48), or as a latent quantity within one common model, such as a bias entry of the state that is zero when the sensor is sound. The second representation means that diagnosis does not always require fitting a separate model for every possible fault. Diagnosis earns its own category for two reasons: the candidates are generated systematically from the plant graph, one for each component or reading that could be at fault, rather than written by hand; and the answer is a ranking. The critical step in either representation is to define the candidate fault events or hypotheses clearly enough that their probabilities can be calculated and compared.

For the hall of Figure 2.1 the family is built on the graph with its reading row kept as a row of its own, $(T_1 - y_T)/\sigma_T$, rather than folded, as in the fit of Section 2.4.1, into the summary row on $g$; a fault must have a row to act on. One candidate is generated for each element that could be at fault, each the extended graph with one entry added: $M_{\text{bias}}$, a bias entry $\Delta_T$ on the reading row, which becomes $(T_1 + \Delta_T - y_T)/\sigma_T$; $M_{\text{path}}$, a shift $\Delta_g$ of the log-conductance away from its prior value, an entry added to both balance rows; and $M_{\text{load}}$, an offset $\Delta_Q$ of the present load $Q_1$ from its nominal value, an entry added to the present-load balance row. Each candidate is fitted to the same readings, its evidence is read from its own $\hat{\mathbf{z}}_j$ and $\boldsymbol{\Sigma}_j$, and the ranking orders them. Which candidates the readings can tell apart is read from the Jacobian before any fit. With the one reading of the hall, each added entry can absorb the misfit of that reading alone, so the three candidates fit it equally well and their ranking rests on their prior rows and prior probabilities. A second reading at another load separates them, because the three added entries then move the two readings in different proportions, which are different columns of $\mathbf{J}$: a bias by equal amounts, a shift of the path in proportion to the loads, and an offset of the present load at the present load only.

### 2.5.4 Expected loss of retuning before the night peak

Questions about a running plant also lead to decisions. The process gain may have drifted since the discharge loop was tuned, and before the night peak the engineer can retune the loop, whose controller's gains lie outside the one-path model of Chapter 1, or leave the tuning, and the probability of an excursion is not enough to decide: the consequence of an excursion and the cost of the action enter as well. Let $e$ be the event that the discharge air leaves its band on the peak's load ramp and overshoots long enough for the IT equipment's protection to act, an excursion, with consequence $C_e = \text{R\$}\,100\,000$ in service credits and recovery, and suppose that if the engineer leaves the tuning its probability is $\Pr(e \mid \mathbf{y}, \text{leave}) = 0.0015$, an illustrative value. An excursion in this sense is a tuning event, not the steady-state exceedance of Question B, the hall crossing the 27.5°C limit, whose probability at 60 kW, under the row of equation (2.29), is 0.280, with 0.283 exact: with $T_c$ pinned, $T_2 = T_c + Q/G$ does not depend on the tuning, so retuning leaves that exceedance where it is, and it does not cure the stiction behind the trip log either. The excursion is a future event that stays random when the state is known, so the loss of leaving the tuning given the state is itself an expectation, the derived quantity

$$L(\text{leave}, \mathbf{z}) = C_e \Pr(e \mid \mathbf{z}, \text{leave}) + 0 \cdot \Pr(\bar{e} \mid \mathbf{z}, \text{leave}) = C_e \Pr(e \mid \mathbf{z}, \text{leave}),$$

with no loss when no excursion occurs, and the expected loss is its location (2.10), an estimation question about a consequence, answered by the expected consequence (2.23) of leaving it. Averaging over the posterior with equation (2.8) turns the conditional probability into its predictive one, and the expected loss is the probability-weighted average of the two outcomes:

$$\mathcal{R}(\text{leave}) = \mathbb{E}[L(\text{leave}, \mathbf{z}) \mid \mathbf{y}] = \Pr(e \mid \mathbf{y}, \text{leave})\, C_e + \Pr(\bar{e} \mid \mathbf{y}, \text{leave}) \cdot 0, \tag{2.50}$$

the second line of equation (2.23) with one event, consequence $C_e$, and no cost of action: probability times consequence. Substituting,

$$\mathcal{R}(\text{leave}) = 0.0015 \times \text{R\$}\,100\,000 + 0.9985 \times 0 = \text{R\$}\,150$$

per night, under these assumptions. It does not mean that the plant will lose exactly R$ 150; the actual loss is either zero or R$ 100 000. It means that R$ 150 is the probability-weighted average of the possible losses, the amount that would be lost per night on average over many nights like this one.

Now suppose retuning costs $c_{\text{retune}} = \text{R\$}\,80$ for the night, in the engineer's slot and recommissioning (a cost, the $c_a$ of equation (2.23)), and lowers the probability of an excursion from 0.0015 to 0.0005. Its loss given the state is $L(\text{retune}, \mathbf{z}) = c_{\text{retune}} + C_e \Pr(e \mid \mathbf{z}, \text{retune})$, and its expected loss is the certain cost of the action plus the expected loss that remains,

$$\mathcal{R}(\text{retune}) = \text{R\$}\,80 + 0.0005 \times \text{R\$}\,100\,000 = \text{R\$}\,130,$$

against $\mathcal{R}(\text{leave}) = \text{R\$}\,150$. The expected net benefit of retuning is $\text{R\$}\,150 - \text{R\$}\,130 = \text{R\$}\,20$ per night, and under these assumptions retuning is the preferred decision. The comparison rests on three numbers, two probabilities and one consequence, and its conclusion changes if any of them moves: the break-even cost of the action is $(0.0015 - 0.0005) \times \text{R\$}\,100\,000 = \text{R\$}\,100$, and, read the other way, retuning and leaving the tuning cost the same when the probability of an excursion under the present tuning is $0.0005 + 80/100\,000 = 0.0013$, the break-even probability. A night whose excursion probability the posterior puts below 0.0013 is a night to leave the tuning. The probabilities here stand for the predictives a posterior would supply; they are given, so the expected losses carry no gap of their own, and a gap in the excursion probability that straddled 0.0013 would be stated with both values and both actions.

### 2.5.5 The general decision formulation

Let $a$ be a control action from a set of candidates, and $L(a, \mathbf{z})$ the loss of taking it when the state is $\mathbf{z}$. The expected loss of the action is its expected consequence (2.23), $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$, the location (2.10) of the derived quantity $L(a, \mathbf{z})$, one for each candidate, and the preferred action is the one that minimizes it:

$$a^* = \arg\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]. \tag{2.51}$$

The retuning example of Section 2.5.4 is this rule with one event, the excursion, and two actions. The change is an important one. The questions of the earlier sections asked what will happen; equation (2.51) asks what should be done, given what might happen. The transition from inference to decisions is what a controller is for.

## 2.6 Measurements, the classification of questions and the derivation workflow

A design question asks whether more information is worth obtaining, and it is answered before the information exists.

### 2.6.1 Uncertainty reduction by a measurement

Suppose the control engineer is uncertain about a quantity because the available record is short, and an additional sensor or a step test could reduce that uncertainty. The question "is the information worth its cost?" is a design question, of the general form (2.16). Its first part is how much the measurement would narrow the posterior: the form with $u$ the posterior covariance, which for linear Gaussian rows is the preposterior covariance (2.17), with the information gain (2.18) as its summary. The plant enters through two objects: the Jacobian $\mathbf{J}_a$ of the candidate rows, which says which entries of the state the measurement acts on and how strongly, and the current covariance $\mathbf{\Sigma}$, which says along which directions the posterior can still shrink.

For the hall, consider $n$ readings of the hall temperature at the present load $Q_1 = 30\,\mathrm{kW}$, each with an error of standard deviation $\sigma_T = 0.5\,\mathrm{K}$, that of the hall sensor of Chapter 1 (independent readings, from sensors at different places of the hall). Each reading is a row on the first entry of the state of equation (2.30),

$$r_{a,j}(\mathbf{z}) = \frac{T_1 - \tilde{y}_j}{\sigma_T}, \qquad \mathbf{J}_{a,j} = \frac{\partial r_{a,j}}{\partial \mathbf{z}} = \begin{bmatrix} 1/\sigma_T & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \end{bmatrix} \mathrm{K}^{-1},$$

so that $n$ readings stack to an $n \times 3$ Jacobian $\mathbf{J}_a$ with $\mathbf{\Lambda}_a = \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a = \mathrm{diag}(4n, 0, 0)$ in $\mathrm{K}^{-2}$. The current $\mathbf{\Sigma}$ is the ridge of equation (2.33), $\sigma_g^2 \mathbf{v}\mathbf{v}^{\mathsf{T}}$, so in the update form (2.27) every product runs through $\mathbf{v}$: $\mathbf{\Sigma}\mathbf{J}_{a,j}^{\mathsf{T}} = \sigma_g^2\,\mathbf{v}\,(2v_1)$ and $\mathbf{J}_{a,j}\mathbf{\Sigma}\mathbf{J}_{a,k}^{\mathsf{T}} = 4\sigma_g^2 v_1^2 = 4\Sigma_{11} = 0.21302$ for every pair of readings. The update therefore removes a multiple of $\mathbf{v}\mathbf{v}^{\mathsf{T}}$ itself, and

$$\mathbf{\Sigma}_a = \frac{\mathbf{\Sigma}}{1 + 0.21302\,n}, \qquad \mathbf{\Sigma}_a = \frac{\mathbf{\Sigma}}{1.21302} = \begin{bmatrix} 0.04390 & 0.08780 & -0.00951 \\ 0.08780 & 0.17561 & -0.01902 \\ -0.00951 & -0.01902 & 0.00206 \end{bmatrix} \quad (n = 1). \tag{2.52}$$

The reading sees only $T_1$, yet every entry of $\mathbf{\Sigma}$ shrinks by the same factor, because the posterior has only the one direction $\mathbf{v}$ to shrink along: one reading removes $1 - 1.21302^{-1/2} = 9.2\,\%$ of the spread of $T_1$, $T_2$ and $g$ alike, so that $\sqrt{\Sigma_{a,33}}$ falls from 0.05 to 0.0454 and $\sqrt{\Sigma_{a,22}}$ from 0.462 to 0.419 K. Table 2.9 gives the result for 0 to 64 readings. The columns are the number of readings, the precision $1/\Sigma_{a,33}$ of $g$, its posterior standard deviation $\sqrt{\Sigma_{a,33}}$, the standard deviation $\sqrt{\Sigma_{a,22}}$ of $T_2$, the fraction of that spread removed, and the expected information gain. The precision of $g$ is $400\,(1 + 0.21302\,n) = 400 + 85.21\,n$ by equation (2.52). The expected information gain is equation (2.18); in the limit both determinants are zero, and the determinant identity $\det\mathbf{\Sigma}/\det\mathbf{\Sigma}_a = \det(\mathbf{I} + \mathbf{J}_a\mathbf{\Sigma}\mathbf{J}_a^{\mathsf{T}})$, with every entry of the $n \times n$ matrix $\mathbf{J}_a\mathbf{\Sigma}\mathbf{J}_a^{\mathsf{T}}$ equal to 0.21302, gives $\mathrm{EIG} = \frac{1}{2}\ln(1 + 0.21302\,n)$. Equation (2.17) evaluated directly at $\sigma_p = 10^{-4}\,\mathrm{K}$, where $\mathbf{\Sigma}^{-1}$ exists, gives the same entries, the precision of 64 readings to within 0.02. The reduction is steep at first and then slows, because the precision grows linearly in $n$ and the spread only as its inverse square root: halving the spread requires a precision of 1600, about fourteen readings, and halving it again a precision of 6400, about seventy.

Table 2.9: Preposterior spread of the hall after $n$ readings at $Q_1 = 30\,\mathrm{kW}$ with $\sigma_T = 0.5\,\mathrm{K}$ (illustrative), from the posterior $\mathbf{\Sigma}$ of equation (2.33) and equation (2.52).

| $n$ [–] | precision $1/\Sigma_{a,33}$ [–] | $\sqrt{\Sigma_{a,33}}$ [–] | $\sqrt{\Sigma_{a,22}}$ [K] | spread removed [–] | EIG [nat] |
|---|---|---|---|---|---|
| 0 | 400.0 | 0.0500 | 0.462 | 0 | 0 |
| 1 | 485.2 | 0.0454 | 0.419 | 9.2 % | 0.097 |
| 4 | 740.8 | 0.0367 | 0.339 | 26.5 % | 0.308 |
| 16 | 1763.3 | 0.0238 | 0.220 | 52.4 % | 0.742 |
| 64 | 5853.3 | 0.0131 | 0.121 | 73.9 % | 1.342 |

### 2.6.2 The value of a step test before deciding

Reducing uncertainty is not necessarily the same as making a better decision. Suppose there are two possible actions, retune the loop and leave the tuning. The best action now minimizes the expected loss, and the expected loss of acting now is

$$\bar{L}_{\mathrm{now}} = \min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}].$$

(The book writes $\bar{L}$ for these minimum expected losses because $C$ names a consequence.) If a candidate test is run first, written $m$ here because $a$ ranges over the actions, with outcome $\tilde{y}_m$, the best action can be chosen again after its value is known, with expected loss $\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}, \tilde{y}_m]$. Before the test is run its result is unknown, so that expected loss is averaged over the possible results, weighted by their predictive distribution given the present evidence:

$$\bar{L}_{\text{after}} = \mathbb{E}_{\tilde{y}_m|\mathbf{y}}\Big[\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}, \tilde{y}_m]\Big] .$$

This is the general form (2.16) for the candidate $m$, with $u$ the minimum expected loss under the updated posterior and the candidate ranked smallest first, and here, unlike for the spread, $u$ depends on the outcome itself and not only on its covariance. The expected value of sample information is the difference,

$$\text{EVSI} = \bar{L}_{\text{now}} - \bar{L}_{\text{after}}, \tag{2.53}$$

the expected improvement in the decision, measured as a reduction in expected loss. It is never negative. For each fixed action, averaging the post-test expected loss over the results returns the present one, $\mathbb{E}_{\tilde{y}_m|\mathbf{y}}\big[\mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}, \tilde{y}_m]\big] = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$, by the law of total expectation; and the minimum over actions inside the average is at most the value of any one action, so $\bar{L}_{\text{after}} \le \min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}] = \bar{L}_{\text{now}}$. Equality holds when no possible result would change the action: a test that cannot change the decision has no value for it, however much it narrows the posterior. If the test costs more than its EVSI, scheduling it is not justified on this criterion.

An illustrative case shows the calculation, with the state here the condition of the process gain, two values, so that every integral is a finite sum, exact, and no second route stands beside it. Suppose the excursion probability 0.0015 of Section 2.5.4 is an average over two conditions of the process gain: drifted from the tuned value, with probability 0.25 and excursion probability 0.0045 under the present tuning, and on it, with probability 0.75 and excursion probability 0.0005; indeed $0.25 \times 0.0045 + 0.75 \times 0.0005 = 0.0015$. Retuning costs R$ 80, restores the match between tuning and gain and so leaves either condition at 0.0005. A step test of the discharge loop, which reads the process gain directly from the response to a valve move, would reveal the condition exactly. Table 2.10 lists, for each condition and for the present state of knowledge, the expected loss of each action and the better one. Its columns are the condition, its probability, the excursion probability under the present tuning, the expected loss of leaving the tuning, that of retuning, and the better action. Without the test the best action is to retune, at R$ 130. With it, a drifted gain leads to retuning at R$ 130 and a gain on its tuned value to leaving the tuning at R$ 50, so

$$\bar{L}_{\text{after}} = 0.25 \times \text{R\$}\,130 + 0.75 \times \text{R\$}\,50 = \text{R\$}\,70, \qquad \text{EVSI} = \text{R\$}\,130 - \text{R\$}\,70 = \text{R\$}\,60.$$

A step test that reveals the condition is worth up to R$ 60 for the night; because it reveals the condition exactly, this is the expected value of perfect information, an upper bound on what any imperfect test could be worth.

This answers Question D. The step test costs R$ 40 of the commissioning slot, a cost the R$ 80 of retuning does not include, since that covers re-commissioning and not the test. The EVSI of R$ 60 for one night exceeds it, a net R$ 60 − R$ 40 = R$ 20, so the test is worth its slot on one night alone. The condition of the gain persists, so later nights add value that this calculation does not count.

Three quantities are now on the table, and they are related but not interchangeable. The reduction of the posterior covariance, equation (2.17), measures how much uncertainty disappears. The expected information gain, equation (2.18), measures how much is learned. The value of information, equation (2.53), measures how much better the resulting decision is expected to be. A reading can score well on the first two and be worth nothing on the third, which is why the choice of the next reading and the choice of an action both need the preposterior but do not ask the same question of it. The value of information is where two

Table 2.10: Expected loss per night of the two actions for a process gain of unknown condition (illustrative): retuning R$ 80 for the night, excursion consequence R$ 100 000.

| Condition | probability [–] | Pr($e$), leave [–] | leave [R$] | retune [R$] | better action |
|---|---|---|---|---|---|
| gain drifted from the tuned value | 0.25 | 0.0045 | 450 | 130 | retune |
| gain on the tuned value | 0.75 | 0.0005 | 50 | 130 | leave the tuning |
| unknown (present) | 1 | 0.0015 | 150 | 130 | retune |

computations meet: the preposterior supplies the outcomes and their weights, and the expected consequence, minimized over the actions, is what it averages. Chapter 23 takes up the choice of a test signal inside a constraint.

### 2.6.3 Identifying the mathematical form of a question

The examples supply a systematic procedure: before computing anything, identify the mathematical objective of the question. Table 2.11 is the reference for that step. Each row is a form of question in the engineer's words, the general form it takes, by equation number, its mathematical form, what it derives to, and its category. These forms recur in the atlas of Appendix A, which collects the questions of each chapter that has a questions table, with the section that answers each.

Many variations of these forms exist. Trip limits on a supervisor, look-ahead over the coming steps, test windows and the transfer of a loop between units each add mathematical structure of their own: joint tails over many entries, integrals of an excursion probability over a horizon, consequences that accrue over a run. The rows of the table, over the general forms of the categories, of a trend and of an expected consequence, already cover a substantial range of inference problems, and a new question is usually classified by recognizing which of them it is, before any computation.

### 2.6.4 The general derivation workflow

The classification is one step of the derivation workflow. Figure 2.6 draws its eight steps: the top row runs from left to right, the bottom row from right to left, and the arrow beneath the bottom row is the return from step 7 to step 3 when a new reading is added. As in the preface, at more length:

**The plant graph (step 0).** The physical system as it is, before any question is asked: its nodes, paths and sources; which quantities are pinned and which are free; the physical relations that must hold among them. The free quantities form the state vector $\mathbf{z}$.

**The question (step 1).** A probability or an expectation of a function of the state, with the unknown on the left of the conditioning bar and everything treated as known on the right.

**The graph extended for the question (step 2).** Scenario nodes for a set point, a unit state or a load not yet applied, candidate faults when something may have changed, and the evidence attached as rows: readings, trip counts, handbook values as priors, each with the modeling assumption it rests on, what that assumption buys and how it can fail.

**The posterior of the state (step 3).** The rows combined into one objective; its minimum, the MAP state; its curvature, the information; and the covariance, derived and then

Table 2.11: Classification of questions by their mathematical form, with the general form of each category from Section 2.3.3 and the forms of a trend and of an expected consequence from Sections 2.3.4 and 2.3.5.

| Control engineer's question | General form | Mathematical form | Derives to | Category |
|---|---|---|---|---|
| What is unknown, and how well is it known? | (2.9), (2.10) | $p(z_i \mid \mathbf{y})$, $\mathbb{E}[z_i \mid \mathbf{y}]$ | a distribution, a mean, a credible interval: $\hat{z}_i$ and $\Sigma_{ii}$ | estimation |
| Does a requirement hold, with the stated assurance? | (2.12) | $\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) \geq \kappa$, $f^*$ the requirement | the posterior integral of the indicator of the requirement; a normal tail under linearization | prediction under a threshold |
| What happens in the coming run? | (2.13) | $\int \Pr(A_{\text{future}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$ | the average over the state, equation (2.41), with the closed form (2.42) beside it where the posterior is a gamma | prediction under a threshold |
| Is it rising, and when does it cross its limit? | (2.21) | $\Pr(\gamma > 0 \mid \mathbf{y})$; $\Pr(t^* \leq s \mid \mathbf{y}) = \Pr(\gamma > 0,\ f_s(\mathbf{z}) \geq f^* \mid \mathbf{y})$ | the posterior read along time, $\mathbf{g}(t)^{\mathsf{T}}\hat{\mathbf{z}}$ and $\mathbf{g}(t)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(t)$; a crossing band by inversion | prediction under a threshold for the sign and a crossing by a date; estimation for a forecast and the distribution of a crossing time |
| Which explanation is correct? | (2.14), (2.15) | $\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\ \Pr(M_j)$ | the evidences and the prior model probabilities; the Bayes factor | hypothesis comparison; diagnosis over a family of faults |
| What is the expected consequence? | (2.23) | $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}] = \int L(a, \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$ | probability times consequence; $L(a, \hat{\mathbf{z}}) + \frac{1}{2}\operatorname{tr}(\mathbf{H}_a \boldsymbol{\Sigma})$ | estimation of a derived quantity |
| What should be done or measured? | (2.23), (2.16) | $\arg\min_a \mathcal{R}(a)$; EVSI | the expected consequence of each action; the preposterior $\boldsymbol{\Sigma}_a$ | design |

computed for the plant, with the exact posterior beside them where the model admits one. It is built once and serves every question.

**The answers (step 4).** Each question read from that posterior, from the engineer's question through its general and specialized forms down to its number: an entry of the state or the gradient of a derived quantity, the quantity at a time asked included; when drawings compete, a ratio of evidences; when actions compete, an expected consequence.

**Evaluation and checks (step 5).** The gaps between the reported values and the exact values or draws collected in a table, the approximations tested, the answers compared with the truth where the record is constructed.

**The report (step 6).** The probability, the Bayes factor or the expected consequence with its band, or a crossing time with its band by inversion, with the exact value beside it where one exists.

**The next reading (step 7).** If the band is too wide, the readings and step tests that could be run, ranked by the width each would remove, or the sensor that most deserves trust, and a return to step 3 with the new row.

The examples of this chapter sit in it as follows. Stating precisely what the engineer needs to know or decide, and defining the events, uncertain quantities, future outcomes, hypotheses or actions that express it, is step 1; adding the full-load scenario node of the hall example, and attaching the readings and priors as rows with the assumptions behind them, is step 2. Step 3 builds the posterior the questions are read from: a fit for the most probable state and its Laplace covariance, a numerical integration, the evidence of each competing model, or another operation appropriate to the problem; for the hall it is the Gauss–Newton fit of three rows, and for the trip log the Newton fit of the log-rate, whose objective is a gamma kernel and so gives the update in closed form as well. Step 4 writes each question in its general form and reduces it, through the specialized form, to posterior marginals, integrals, gradients, evidences or decision objectives, and evaluates them. Step 5 checks assumptions, identifiability, approximations, numerical convergence and the sensitivity of the answer to modeling choices, as the tolerance, exact-value and sampling checks of the hall example did. Step 6 communicates the answer, its uncertainty, its limitations and what it means for the run, and step 7 is the design question of Section 2.6.1.

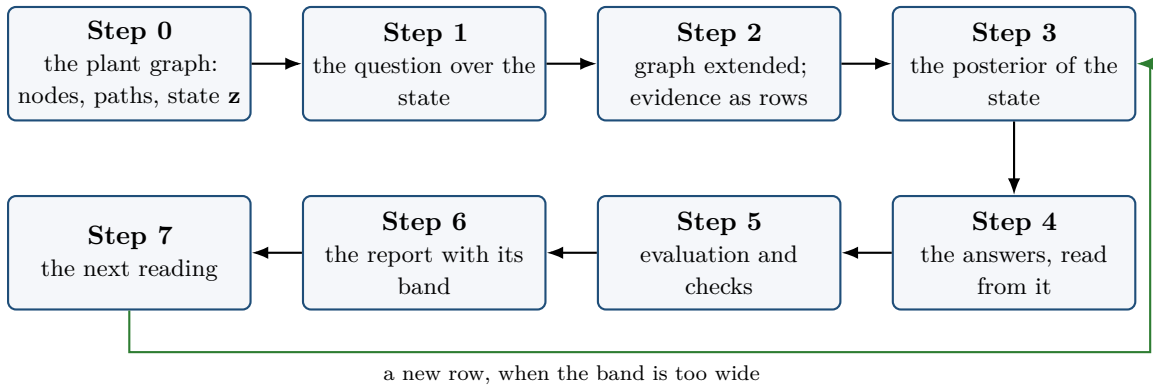



Figure 2.6: The derivation workflow of the book. Step 0 is built once per plant; steps 1 to 7 are one pass; step 3 is shared by every question put to the same plant and evidence; the green return arrow adds the reading chosen at step 7 as a new row and rebuilds the posterior from step 3.

The workflow is broader than any particular software implementation. The derivation fixes what has to be calculated before a computational tool is chosen, and the tool then calculates it.

**Principle 2.3.** *The derivation tells the computational tool what must be calculated. A computational tool should not dictate which control question is asked.*

### 2.6.5 Common mistakes in formulating control questions

Six mistakes of formulation recur on a commissioning sheet, and each is a confusion between two questions that look alike.

**Common mistake**

Confusing uncertainty about a trip rate with the probability of a clean run. For a loop with an unknown trip rate $\lambda$, the probability of no trip over a horizon $\tau$ at a given rate is $\mathrm{e}^{-\lambda\tau}$. The question $\mathbb{E}[\mathrm{e}^{-\lambda\tau} \mid \mathbf{y}]$ is the posterior predictive probability of a clean horizon: how likely the next run is to pass without a trip. The question $\Pr(\mathrm{e}^{-\lambda\tau} > 0.99 \mid \mathbf{y})$ is the posterior probability that the loop's underlying rate meets a target written as a 99 % chance of a clean horizon: how much support there is for the loop meeting its target. They are different numbers and answer different questions.

**Common mistake**

Confusing prediction with diagnosis. When the hall runs warm, the prediction $\Pr(T > T_{\max} \mid \mathbf{y})$ concerns a temperature event, and the diagnosis $\Pr(M_{\text{fouled}} \mid \mathbf{y})$ concerns an explanation, a fouled coil against a sound one. They can use the same readings but require different calculations: a tail of a posterior in one model against a comparison of evidences across models.

**Common mistake**

Ignoring the time horizon. $\Pr(T > 12\,\mathrm{h} \mid \mathbf{y})$ and $\Pr(T > 8760\,\mathrm{h} \mid \mathbf{y})$ are not interchangeable statements about the loop; for the trip log of Section 2.4.4 the first, a clean night run, is 0.9646 (0.9686 exact) and the second, a clean year, is 0.00004 (exact, $(1500/10\,260)^4 = 0.0005$). A probability of a clean horizon is incomplete without its horizon and its operating conditions.

**Common mistake**

Forgetting dependence. In general $\Pr(A \cap B) \neq \Pr(A)\,\Pr(B)$; the product is justified only under independence. A shared pump, a shared supply-air set point, the weather and a night's test work create dependence between loops and units, and Table 2.8 shows that a small common cause can multiply the probability of two loops tripping in one night several times over.

**Common mistake**

Treating the model assumptions as certain. A calculated probability may be numerically precise while depending strongly on uncertain modeling assumptions. A temperature-exceedance probability computed under the steady-state model of equation (2.2) may be unsuitable for the minutes after the chilled water stops, when the hall air rises through a transient that the steady-state model does not describe and the supply air is no longer held at $T_c$. Every probability is conditional on its model, $\Pr(A \mid \mathbf{y}, M)$, and must be interpreted in that context.

**Common mistake**

Reading a flat loop trend as a healthy plant. The trend shows the reading held at its set point: the discharge reading lies flat at 18°C. That is a statement about the loop, and it carries no band on the hall. With that same flat trend, the fit of Chapter 1 gives the hall $23.07 \pm 0.481$ K, the conductance $5.917 \pm 0.561$ kW/K and a probability of only 0.246 (0.281 exact) that the hall holds 27.5°C at 60 kW.

The trip logs, the readings and the two loops of this chapter are constructed, illustrative records, not a plant's measurements. The questions of this chapter were all written with conditional probabilities and the posterior was taken as given; the next chapter, Chapter 3, derives conditional probability and Bayes' rule, separates aleatory variability from epistemic uncertainty, and shows how readings change the uncertainty about the state of the hall through marginalization, dependence and conditional independence on the plant graph.

## 2.7 What to retain

**Definition 2.1** (Control question)**.** A control question about a plant is a probability or an expectation of a function of the state of its plant graph, written with the unknown on the left of the conditioning bar and the evidence and the model on the right, $\Pr(A \mid \mathbf{y}, M) = \int \mathbf{1}_A(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}$ for a present event and $\int \Pr(A_{\text{future}} \mid \mathbf{z}, M)\, p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}$ for a future one. Its form decides its category, estimation, prediction under a threshold, hypothesis comparison, diagnosis or design, each with its general forms, equations (2.9) to (2.16); the question decides which of five computations answers it, the posterior of the state, the trend, the evidence, the preposterior or the expected consequence (Table 2.3), two of which carry general forms of their own, the trend form (2.21) and the expected-consequence form (2.23); and the plant enters only in the specialized form, through the entry or function of the state, the event or family of models, and the objects of the fit that carry the answer.

Table 2.12 lists the chapter's terms with their expressions, units and the subsection that defines each, grouped by their role.

Table 2.12: The terms of control questions and where each is defined.

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| Event | $A$, e.g. $T > 28$°C | – | a statement true or false | 2.2.1 |
| Conditional probability | $\Pr(A \mid \mathbf{y}, M)$ | – | probability given evidence and model | 2.2.3 |
| Event probability from a density | $\int_5^7 p(G \mid \mathbf{y})\, \mathrm{d}G$ | – | probability of the range $5 < G < 7$ | 2.2.4 |
| Probability over the state | $\int \mathbf{1}_A(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$ | – | posterior expectation of the indicator | 2.3.1 |
| Posterior predictive | $\int \Pr(A_{\text{future}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$ | – | future event averaged over the state | 2.3.2 |
| *General forms* | | | | |
| Estimation | $\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$; $\mathbb{E}[f \mid \mathbf{y}]$, $\mathrm{Var}(f \mid \mathbf{y})$ | – | posterior of a derived quantity | 2.3.3 |
| Credible interval | $\Pr(f \le f_L \mid \mathbf{y}) = 0.025$, $\Pr(f \le f_U \mid \mathbf{y}) = 0.975$ | – | central 95 % band; quantiles map through monotone $f$ | 2.3.3 |

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| Threshold and assurance | $\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) \geq \kappa$ | – | probability that the requirement $f \geq f^*$ holds, against $\kappa$ | 2.3.3 |
| Predictive | $\int p(y_{\text{new}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$ | – | future observation averaged over the state | 2.3.3 |
| Comparison | $B_{AB} = p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$ | – | evidence ratio of two drawings | 2.3.3 |
| Diagnosis | $\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\ \Pr(M_j)$, ranked | – | comparison over a family of faults | 2.3.3 |
| Design | $U(a) = \mathbb{E}_{\tilde{y}_a \mid \mathbf{y}}[u(p(\mathbf{z} \mid \mathbf{y}, \tilde{y}_a))]$ | varies | preposterior expectation | 2.3.3 |
| Preposterior covariance | $\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$ | varies | spread after a future measurement | 2.3.3 |
| Expected information gain | $\frac{1}{2}\ln(\det \boldsymbol{\Sigma}/\det \boldsymbol{\Sigma}_a)$ | nat | what a measurement teaches | 2.3.3 |
| Trend | $\Pr(\gamma > 0 \mid \mathbf{y})$; $\Pr(t^* \leq s \mid \mathbf{y}) = \Pr(\gamma > 0,\ f_s(\mathbf{z}) \geq f^* \mid \mathbf{y})$; $p(f_t(\mathbf{z}) \mid \mathbf{y})$; read with $\mathbf{g}(t)^{\mathsf{T}}\hat{\mathbf{z}}$, $\mathbf{g}(t)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(t)$ | – | sign, crossing and forecast: the posterior read along time | 2.3.4 |
| Rows of a record in windows | $\mathbf{J} = \mathbf{D}\mathbf{X}$, $[\boldsymbol{\Lambda}]_{\gamma\gamma} = \sum_k d_k^2 (t_k - \bar{t})^2$ | varies | a window's distance enters squared, windows add linearly | 2.3.4 |
| Expected consequence | $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$, $a^* = \arg\min_a \mathcal{R}(a)$; spread $\sqrt{\mathbf{g}_a^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}_a}$ | varies | probability times consequence, minimized over actions | 2.3.5 |
| Curvature term | $L(a, \hat{\mathbf{z}}) + \frac{1}{2}\operatorname{tr}(\mathbf{H}_a\boldsymbol{\Sigma})$ | varies | what the plug-in loss misses, in the state's coordinates | 2.3.5 |
| *Computations* | | | | |
| Posterior | $p(\mathbf{z} \mid \mathbf{y}) \propto \mathrm{e}^{-\mathcal{L}(\mathbf{z})}$; $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ | varies | what estimation and prediction integrate | 2.3.6 |
| Evidence | $p(\mathbf{y} \mid M) = \int p(\mathbf{y} \mid \mathbf{z}, M)\, p(\mathbf{z} \mid M)\, \mathrm{d}\mathbf{z}$ | – | normalizing constant of a model | 2.3.6 |
| Preposterior | $p(\mathbf{z} \mid \mathbf{y}, \tilde{y}_a)$ averaged over $p(\tilde{y}_a \mid \mathbf{y})$ | varies | the expected posterior under a measurement not yet taken | 2.3.6 |
| Information adds | $\mathbf{J}^{\mathsf{T}}\mathbf{J} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a$ | varies | stacked rows of a candidate | 2.3.6 |
| Update form | $\boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{J}_a^{\mathsf{T}}(\mathbf{I} + \mathbf{J}_a\boldsymbol{\Sigma}\mathbf{J}_a^{\mathsf{T}})^{-1}\mathbf{J}_a\boldsymbol{\Sigma}$ | varies | $\boldsymbol{\Sigma}_a$ when $\boldsymbol{\Sigma}$ is singular | 2.3.6 |
| *Building blocks of the fit* | | | | |
| State | $\mathbf{z} = (T_1, T_2, g)^{\mathsf{T}}$ | varies | every free quantity of the graph at once | 2.4.1 |
| Rows and objective | $r_k(\mathbf{z})/\sigma_k$; $\mathcal{L} = \frac{1}{2}r^{\mathsf{T}}r$ | – | physics, reading and prior rows in standard deviations | 2.4.1 |
| Jacobian | $\mathbf{J} = \partial r/\partial \mathbf{z}$ | varies | sensitivity of each row to each entry | 2.4.1 |
| Normal equations | $(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}r$ | varies | Gauss–Newton step to $\hat{\mathbf{z}}$ | 2.4.1 |
| Laplace covariance | $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ | varies | band of every entry of the state | 2.4.1 |
| Ridge of exact rows | $\boldsymbol{\Sigma} \to \sigma_g^2\mathbf{v}\mathbf{v}^{\mathsf{T}}$ as $\sigma_p \to 0$ | varies | entries tied by the physics move as one | 2.4.1 |

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| Band of a derived quantity | $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$, $\mathbf{g} = \partial f/\partial \mathbf{z}$ | varies | delta method | 2.4.1 |
| *Specialized forms and answers* | | | | |
| Requirement probability of the hall | $\Phi\big((T_{\max} - \hat{z}_2)/\sqrt{\Sigma_{22}}\big) = 0.720$ | – | reported, the normal tail of the entry | 2.4.2 |
| Its exact value | $\Phi\big((\hat{z}_3 - \ln(G^*/G_{\mathrm{ref}}))/\sqrt{\Sigma_{33}}\big) = 0.717$ | – | the bound on $T_2$ mapped along the ridge to the third entry; gap 0.003 | 2.4.2 |
| Trip rate | $\hat{\lambda} \pm \hat{\lambda}\,\sigma_\omega = 0.002667 \pm 0.001333$ | $\mathrm{h}^{-1}$ | reported, from the log-rate | 2.4.3 |
| Exact posterior of the rate | $\mathrm{Gamma}(a + k,\ b + T_{\mathrm{tot}})$ | –, h | the same trip log, conjugate | 2.4.3 |
| Clean run | $\sum_j w_j\, \mathrm{e}^{-\lambda_{\mathrm{ref}} \mathrm{e}^{\omega_j} \tau} = 0.9646$ | – | reported, no trip averaged over the normal of the log-rate | 2.4.4 |
| Its exact value | $(b'/(b' + \tau))^{a'} = 0.9686$ | – | the same average over the gamma posterior; gap 0.00406 | 2.4.4 |
| Plug-in clean run | $\mathrm{e}^{-\mathbb{E}[\lambda \mid \mathbf{y}]\tau}$ | – | the same at a point estimate; below the exact value | 2.4.5 |
| Both loops clean | $1 - \Pr(A_A) - \Pr(A_B) + \Pr(A_A \cap A_B)$ | – | inclusion and exclusion | 2.4.6 |
| Both loops tripping | $\Pr(A_A \mid \mathbf{y})\ \Pr(A_B \mid A_A, \mathbf{y})$ | – | intersection; product rule; a shared pump | 2.4.7 |
| Model posterior | $\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\ \Pr(M_j)$ | – | probability of a model | 2.5.1 |
| Bayes factor of the air path | $B_{AB} = p(\mathbf{y} \mid M_1)/p(\mathbf{y} \mid M_0)$, $M_A = M_1$, $M_B = M_0$ | – | relative support from the evidence | 2.5.1 |
| Expected loss of an action | $\mathcal{R}(a) = c_a + C_e\ \Pr(e \mid \mathbf{y}, a)$ | R$ | retuning against leaving the tuning | 2.5.4 |
| Break-even probability | $\Pr(e \mid \mathbf{y}, \mathrm{retune}) + c_{\mathrm{retune}}/C_e$ | – | the probability at which retuning and leaving the tuning cost the same | 2.5.4 |
| Optimal action | $a^* = \arg\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$ | – | decision by expected loss | 2.5.5 |
| Readings on the hall | $\boldsymbol{\Sigma}_a = \boldsymbol{\Sigma}/(1 + 0.21302\, n)$ | varies | every entry shrinks along the ridge | 2.6.1 |
| Value of information | $\mathrm{EVSI} = \bar{L}_{\mathrm{now}} - \bar{L}_{\mathrm{after}}$ | R$ | what a step test improves | 2.6.2 |
| *The book's questions* | | | | |
| Control questions | twenty-six lead questions | – | general form, category, computation and chapter of each | 2.3.8 |

## 2.8 Exercises

**Exercise 2.1** (Formulation)**.** For each statement, name the category of the question, the expression that answers it and the computation beneath it. (a) What is the probability that the loop runs clean through the next 12 h? ($\Pr(T > 12\,\mathrm{h} \mid \mathbf{y})$, a posterior predictive probability over the stated horizon; prediction under a threshold; the posterior.) (b) What is the trip

rate of the loop, given its trip log? ($p(\lambda \mid \mathbf{y})$ or a summary of it; estimation; the posterior.) (c) Two loops tripped in the same night; is a common cause, the pump, more likely than two independent causes? ($\Pr(M_{\text{common}} \mid \mathbf{y})$ against $\Pr(M_{\text{independent}} \mid \mathbf{y})$, or their Bayes factor; hypothesis comparison; the evidence.) (d) At which hour of the evening should the loop be retuned to minimize the expected cost of the night? ($\arg\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$ over the candidate hours; a decision, in the design category; the expected consequence.) (e) Is the trip rate of a loop rising, and in which period will it pass the site's valve-replacement limit? ($\Pr(\gamma > 0 \mid \mathbf{y})$ and $\Pr(t^* \le s \mid \mathbf{y})$ over a state that carries the log-rate and its trend; prediction under a threshold; the trend.) (f) Of the questions A to D of Table 2.1, which two are answered from the same posterior without a new calculation of it? (A and B: an estimation and a threshold probability read from one posterior.)

**Exercise 2.2** (Formulation)**.** A loop has a constant trip rate $\lambda$ with a known posterior, and the site's target for a valve in service, not its valve-replacement limit of Exercise 2.5, asks for at least a 99 % probability of a clean 24 h day, two runs. (a) Formulate the posterior predictive probability of a clean day and the probability that the loop's underlying rate meets the target. ($\mathbb{E}[\mathrm{e}^{-24\lambda} \mid \mathbf{y}] = \int_0^\infty \mathrm{e}^{-24\lambda} p(\lambda \mid \mathbf{y})\,\mathrm{d}\lambda$; $\Pr(\mathrm{e}^{-24\lambda} \ge 0.99 \mid \mathbf{y}) = \Pr(\lambda \le -\ln 0.99/24 \mid \mathbf{y}) = \Pr(\lambda \le 4.188 \times 10^{-4}\,\mathrm{h}^{-1} \mid \mathbf{y})$.) (b) Evaluate both from the posterior of the state, $\hat{\omega} = 0.98083$ and $\sigma_\omega = 0.5$, and set beside each the exact value from the posterior Gamma$(4, 1500\,\mathrm{h})$ of the loop's trips, using for the second the integer-shape identity $\Pr(\lambda \le x \mid \mathbf{y}) = 1 - \sum_{j=0}^{a'-1} \mathrm{e}^{-b'x}(b'x)^j/j!$. (From the state, equation (2.41) with $\tau = 24$ h gives 0.9307, and $\Phi\big((\ln 0.41876 - 0.98083)/0.5\big) = \Phi(-3.703) = 0.0001$. Exactly, $(1500/1524)^4 = 0.9385$, a gap of 0.0078; $b'x = 0.6281$ and the four Poisson terms 0.5336, 0.3352, 0.1053 and 0.0220 leave 0.0039, a gap of 0.0038.) (c) Explain why they differ. (The first averages the probability of no trip over the uncertainty in the rate and asks how likely tomorrow is to be clean; the second asks how much posterior probability supports the target on the rate itself. Here a clean day is likely, 0.93 and 0.94 exact, while the target is almost certainly not met by either route, because the posterior puts almost no mass below $4.2 \times 10^{-4}\,\mathrm{h}^{-1}$.)

**Exercise 2.3** (Derivation)**.** The hall obeys $T = T_c + Q/G$ with $T_c = 18$°C and $Q = 60$ kW. (a) With the state, rows and $\mathbf{\Sigma}$ of equations (2.30) to (2.33), find the probability that the full-load temperature remains below 28°C from the normal of the entry $T_2$, then the exact value from the decreasing map between $G$ and $T_2$ along the ridge. ($\Pr(T_2 < 28\text{°C} \mid \mathbf{y}) \approx \Phi\big((28 - \hat{z}_2)/\sqrt{\Sigma_{22}}\big) = \Phi\big((28 - 27.231)/0.4615\big) = \Phi(1.667) = 0.952$. Exactly, along $\mathbf{v}$ the bound on $T_2$ is $G > Q_2/(28 - T_c) = 6$ kW/K, that is $g > \ln 6$ on the third entry, so $\Pr(T_2 < 28\text{°C} \mid \mathbf{y}) = \int_6^\infty p(G \mid \mathbf{y})\,\mathrm{d}G = \Phi\big((\hat{z}_3 - \ln 6)/\sqrt{\Sigma_{33}}\big) = \Phi(1.601) = 0.945$; both are rows of Table 2.6. At $\kappa = 0.95$ the first accepts and the second does not, so both are stated.) (b) Now let the supply air be uncertain as well, as it is when the discharge valve is near its stop, so that $G$ and $T_c$ have a joint posterior $p(G, T_c \mid \mathbf{y})$, and derive an integral for $\Pr(T > T_{\max} \mid \mathbf{y})$ without assuming that $G$ and $T_c$ are independent. (The event is $T_c + Q/G > T_{\max}$, so $\Pr(T > T_{\max} \mid \mathbf{y}) = \iint \mathbf{1}\big[T_c + Q/G > T_{\max}\big]\, p(G, T_c \mid \mathbf{y})\,\mathrm{d}G\,\mathrm{d}T_c$; the supply air becomes an entry of the state, and the joint density is required because the readings can leave it and the conductance correlated, a nonzero off-diagonal entry of $\mathbf{\Sigma}$.)

**Exercise 2.4** (Derivation)**.** Two loops on one site have probabilities $\Pr(A_A) = 0.08$ and $\Pr(A_B) = 0.04$ of tripping during the night. (a) Assuming independence, derive the probability that both trip and the probability of a clean night on both. ($\Pr(A_A \cap A_B) = 0.08 \times 0.04 = 0.0032$; $\Pr(\bar{A}_A \cap \bar{A}_B) = 0.92 \times 0.96 = 0.8832$.) (b) Now let both valves sit behind one chilled-water pump that stops with probability $q_S = 0.01$ in the night, with the marginals unchanged. Derive the joint probability from equation (2.47), the conditional probability $\Pr(A_B \mid A_A)$ and the probability of a clean night on both. ($q_A = 0.07/0.99 = 0.07071$, $q_B = 0.03/0.99 = 0.03030$; $\Pr(A_A \cap A_B) = 0.01 + 0.99 \times 0.07071 \times 0.03030 = 0.01212$; $\Pr(A_B \mid A_A) = 0.01212/0.08 = 0.152$;

$\Pr(\bar{A}_A \cap \bar{A}_B) = 1 - 0.12 + 0.01212 = 0.89212$; the probability of both tripping is 3.8 times the independent value.)

**Exercise 2.5** (Derivation)**.** The trips of a sister loop's older valve are counted in six periods of forty days, period $j$ centered at $t = j$ so that the log ends at $t = 6.5$, the first period with 36 logged days, and number 5, 6, 7, 9, 11 and 14. The log trip rate, in trips per logged day, is carried as a straight line in time, $f_t(\mathbf{z}) = \alpha + \gamma(t - \bar{t})$ with $t$ in periods and $\bar{t} = 3.5$, and the fit of the six count rows with flat priors gives $\hat{\mathbf{z}} = (-1.5768,\ 0.1989)^{\mathsf{T}}$ and $\boldsymbol{\Sigma}$ with entries $\Sigma_{\alpha\alpha} = 0.021811$, $\Sigma_{\alpha\gamma} = -0.004328$ and $\Sigma_{\gamma\gamma} = 0.007259$. The site's valve-replacement limit is 0.5 trips a day, $f^* = \ln 0.5$; this older valve, at 0.139 trips a day already in its first period, is well past the target of Exercise 2.2. (a) From the trend form (2.21) under the Laplace posterior, find the probability that the rate is rising. ($\Phi(0.1989/\sqrt{0.007259}) = \Phi(2.335) = 0.990$.) (b) Find the crossing time at the MAP state and say why it stands at the median of the crossing time at every printed digit, though not exactly. ($t^*(\hat{\mathbf{z}}) = 3.5 + (\ln 0.5 + 1.5768)/0.1989 = 7.94$ periods, $(7.94 - 6.5) \times 40 = 58$ days after the log ends; the second argument of $\Phi_2$ in equation (2.22) is zero at $s = t^*(\hat{\mathbf{z}})$, so the threshold line is one half there, and the first passage subtracts from it the states that stand at or above the limit with $\gamma \le 0$, a mass of $2.8 \times 10^{-11}$, so the median lies about $1.3 \times 10^{-10}$ periods later.) (c) Find the probability that the rate has crossed the limit by $t = 8$. ($\mathbf{g}(8) = (1,\ 4.5)^{\mathsf{T}}$, $\mathbf{g}(8)^{\mathsf{T}}\hat{\mathbf{z}} = -0.6816$, $\mathbf{g}(8)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(8) = 0.021811 - 9 \times 0.004328 + 20.25 \times 0.007259 = 0.1299$, and the covariance of $\gamma$ and $f_8$ is $\Sigma_{\alpha\gamma} + 4.5\,\Sigma_{\gamma\gamma} = 0.02834$, a correlation of 0.923. The threshold line is $\Phi\big((-0.6816 + 0.6931)/0.3604\big) = \Phi(0.032) = 0.513$, and the first passage is $\Pr(\gamma > 0,\ f_8 \ge f^* \mid \mathbf{y}) = \Phi_2(2.335,\ 0.032;\ 0.923) = 0.513$, less by $2.8 \times 10^{-11}$.) (d) Explain why the crossing time has no posterior mean. ($t^* = \bar{t} + (f^* - \alpha)/\gamma$ divides by the trend entry, whose Gaussian posterior has a positive density at $\gamma = 0$, and the integral of $1/\gamma$ against such a density diverges.)

**Exercise 2.6** (Application)**.** The probability of an excursion at tonight's peak is 0.004 if the tuning is left as it is, and an excursion costs R$ 250 000. Retuning the hall's loop for the night costs R$ 400 and lowers the probability to 0.001. (a) Derive the expected cost of each action, and state which action minimizes it. (The expected consequence (2.23) of each action, probability times consequence plus the certain cost. Leaving the tuning: $0.004 \times$ R$ 250 000 = R$ 1000; retuning: R$ 400 + $0.001 \times$ R$ 250 000 = R$ 650; retuning minimizes the expected cost, with an expected saving of R$ 350.) (b) Find the cost of the action at which the decision would reverse. ($(0.004 - 0.001) \times$ R$ 250 000 = R$ 750.)

**Exercise 2.7** (Application)**.** Two models are competing explanations of a hall sensor that reads warmer than the model of the hall expects, $M_0$ a sound sensor and $M_1$ a biased one, with $\Pr(M_0) = \Pr(M_1) = 0.5$ and evidences $p(\mathbf{y} \mid M_0) = 0.02$ and $p(\mathbf{y} \mid M_1) = 0.08$. (a) Derive the Bayes factor $B_{10}$ and the posterior probability $\Pr(M_1 \mid \mathbf{y})$. ($B_{10} = 0.08/0.02 = 4$, $\ln B_{10} = 1.386$ nats; $\Pr(M_1 \mid \mathbf{y}) = 0.08 \times 0.5/(0.08 \times 0.5 + 0.02 \times 0.5) = 0.80$, assuming that the two models exhaust the possibilities under consideration.) (b) Repeat with $\Pr(M_1) = 0.2$, and say what the comparison of the two answers shows. ($0.08 \times 0.2/(0.08 \times 0.2 + 0.02 \times 0.8) = 0.50$; the Bayes factor is the same 4 in both cases, because it depends only on the evidence, while the posterior model probability also depends on the prior.)

# Chapter 3

# Uncertainty and Conditional Probability

## 3.1 Uncertainty and evidence in a control question

The discharge trend of the hall's cooling unit sits at 18.0°C with small scatter, and the loop's trip flag, which fires when the loop's error stays outside a band past an on-delay and the loop goes to manual, has fired twice in the six weeks of its log. Read from the trend, the conclusion is quick: the loop is tight, because the reading does not leave its set point, and the valve is sticking, because the flag has said so twice. The engineer also knows that a second loop, loop B, hangs on the same chilled-water pump, and wonders whether loop B needs the same attention.

The trend does not settle either point, and the feedback is the reason. A controller holds its reading, so the trend is a statement about the loop before it is one about the plant, and three things stay open. How much of the scatter on the trend is the sensor is open. What a flag means is open: the flag fires for more reasons than stiction, and valves really stick rarely. What the two loops share on one pump is open too. Four situations of this kind recur through the chapter. A hall reading scatters about the air temperature the held supply air leaves, and the probability that the next reading crosses the hall's 24°C limit is wanted. The trip flag has fired, and the probability that the valve is really sticking is wanted. Two loops draw chilled water through one pump, and the probability that both trip in the same run is wanted. The hall has one reading, and the probability that the hall air exceeds its limit is wanted. Chapter 1 represented the hall as a plant graph whose free quantities form the state vector $\mathbf{z}$, and Chapter 2 wrote the engineer's questions as probabilities over that state, with the unknown on the left of the conditioning bar and the evidence on the right. The question of this chapter is the mechanism between the two:

*how new evidence changes what is known about a running loop, and how dependencies between loops and between physical quantities affect a probability calculation.*

**Jargon: aleatory and epistemic uncertainty, conditional probability, prior, likelihood, evidence, posterior, marginalization, nuisance quantity, conditional independence, common cause**

*Aleatory* uncertainty is the variability of outcomes that the model treats as random even when its parameters are known, such as the reading a sensor will show in the next window when the hall air is known. *Epistemic* uncertainty is incomplete knowledge of quantities that have one true value, such as the conductance of the hall's air path or the calibration offset of a sensor; more evidence can reduce it. A *conditional probability* $\Pr(A \mid B)$ is the probability of event $A$ among the situations in which event $B$ occurs. In Bayes' theorem the *prior* is the probability of a hypothesis before an observation, the *likelihood* is the probability of the observation under the hypothesis, the *evidence* (also *marginal likelihood*) is the probability of the observation averaged over the hypotheses, and the *posterior* is the probability of the hypothesis after the observation. *Marginalization* removes an unknown from a joint distribution by summing or integrating over it; an unknown removed this way because the question does not ask about it is a *nuisance quantity*. Two quantities are *conditionally independent* given a third when, once the third is known, learning one says nothing more about the other. A *common cause* is a shared condition, such as a sag in the flow of a chilled-water pump behind two valves, that raises the probability of a trip in several loops at once.

**Key idea**

Conditioning on evidence restricts attention to the situations consistent with it and renormalizes; Bayes' theorem does this for a hypothesis, marginalization averages over every unknown the question does not name, and the dependence among unknowns, through a common cause or a shared physical relation, is what lets one observation inform another quantity and what an independence assumption silently discards.

A loop trend leaves two kinds of not knowing: what the next window will show, and what the plant's gains are. Keeping them apart changes the probability that a reading crosses a limit, and each kind has its own distribution.

### 3.1.1 Aleatory scatter of a controlled reading and of a load

A controller holds a reading about its set point, and a trend shows the scatter that remains. This part of the book treats the discharge controller as ideal: with a true discharge sensor, the supply air sits at its set point $T_c = 18°\mathrm{C}$, which fixes the hall air for a given load and conductance. The reading computed here is the hall sensor's, which no controller holds; the loop's own reading returns at the end of this subsection. Suppose the hall air were known exactly: take the estimate of Chapter 1, $T = 23.07°\mathrm{C}$, as if it were the truth. Even then the reading of the next five-minute window cannot be predicted. Let $y_T$ be the reading of one window. The sensor adds an error that changes from window to window, $y_T \mid T \sim \mathcal{N}(T, \sigma_T^2)$ with $\sigma_T = 0.5\,\mathrm{K}$, so the probability that the reading of a window shows above a level $c$ is the integral of the normal density from $c$ to infinity,

$$\Pr(y_T > c \mid T) = \int_c^\infty \frac{1}{\sigma_T}\,\varphi\!\left(\frac{s - T}{\sigma_T}\right) \mathrm{d}s = 1 - \Phi\!\left(\frac{c - T}{\sigma_T}\right), \tag{3.1}$$

where $s$ is the variable of integration and $\varphi$ and $\Phi$ are the standard normal density and cumulative distribution function. For the 24°C limit,

$$\Pr(y_T > 24°\mathrm{C} \mid T = 23.07°\mathrm{C}) = 1 - \Phi\!\left(\frac{24 - 23.0701}{0.5}\right) = 1 - \Phi(1.8598) = 0.0315.$$

With the hall air perfectly known there is still a probability 0.0315, one window in 32, that the reading shows the hall above its limit; over a twelve-hour run of 144 windows a trend that colors a reading red above 24°C would color about 4.5 windows red with the hall inside its limit throughout. This uncertainty arises from the variability that the reading process itself represents, and it is called *aleatory*. The IT load is a second source of it: server utilization moves the load from window to window, and a scatter of 1.5 kW about 30 kW (5 %, illustrative) moves the hall air itself by $1.5/5.917 = 0.2535$ K at the estimated conductance 5.917 kW/K. The reading then scatters by $\sqrt{0.5^2 + 0.2535^2} = 0.561$ K and crosses 24°C in 0.0486 of the windows. The rest of the chapter holds the load fixed within a window and keeps the sensor's scatter alone.

Figure 3.1(a) shows it. The horizontal axis numbers twelve windows of the same hall with the same known air temperature, and the vertical axis is the reading of each. The dots are the twelve quantiles of the normal distribution $\mathcal{N}(23.07, 0.5^2)$, placed in an arbitrary order; the dashed line marks the true hall air and the red line the limit. Identical windows of an identical, exactly known hall read from 22.20 to 23.94°C: knowing the hall fixes the distribution of the reading, not the reading. The discharge reading about its set point is the same model with that sensor's own spread: a loop that holds its reading has not removed the scatter on the next one.

### 3.1.2 Epistemic uncertainty about the plant's gains

Now suppose the hall air is not known, because what fixes it at a given load, the effective conductance $G$ of the air path to the supply air, is not; $1/G$ is the process gain from the load to the hall air. The handbook gives 7.5 kW/K, right to a factor of about 1.35. What is known about the conductance is then a probability distribution given the evidence, $p(G \mid \mathbf{y})$, or equivalently $p(g \mid \mathbf{y})$ for the log-conductance $g = \ln(G/G_{\text{ref}})$ with $G_{\text{ref}} = 6$ kW/K. The conductance might plausibly be $7.5/1.35 = 5.556$, 7.5 or $7.5 \times 1.35 = 10.125$ kW/K, with different probabilities, or a density, assigned to those values. This is *epistemic* uncertainty: uncertainty caused by incomplete knowledge of the characteristics of the plant, which additional observations can reduce. The calibration of a sensor is the other common case. A sensor that reads off by a fixed, unknown offset reads off in every window; the offset has one true value, and the evidence that reduces what is known about it is a comparison with an independent measurement, not more windows of the same sensor.

Figure 3.1(b) draws such a distribution. The horizontal axis is the conductance in kW/K and the vertical axis the prior density. The curve is the handbook prior of Chapter 1, $g \sim \mathcal{N}(\ln(7.5/6), 0.3^2)$, drawn as the lognormal density of $G$; the dashed line is its median, 7.5 kW/K, and the three dots are the three plausible values named above. The conductance is a single number; the spread of the curve is the spread of knowledge about it.

### 3.1.3 The distinction on the hall

In a running plant the distinction can depend on what information is available and on how the model is defined. A future load may be modeled as random, aleatory, while the load at the present moment is a fixed but poorly metered quantity, epistemic. The distinction is therefore a statement about the model, and it changes the calculation.

If the hall air is known, the probability that the next reading shows above the limit is equation (3.1). If the hall air is uncertain, the probability that the next reading $y_{\text{new}}$ crosses $c$ is the conditional probability averaged over what is known about the hall, which is the law of total probability derived in Section 3.4.2:

$$\Pr(y_{\text{new}} > c \mid \mathbf{y}) = \int \Pr(y_{\text{new}} > c \mid T)\, p(T \mid \mathbf{y})\, \mathrm{d}T = \int \left[1 - \Phi\!\left(\frac{c - T}{\sigma_T}\right)\right] p(T \mid \mathbf{y})\, \mathrm{d}T. \quad (3.2)$$

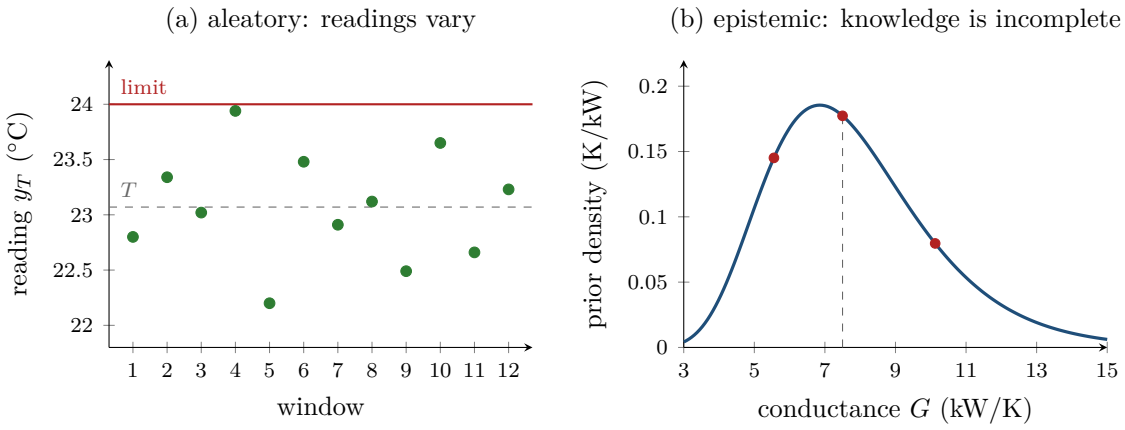



Figure 3.1: The two kinds of uncertainty. (a) Readings of twelve windows of a hall whose air is known, $T = 23.07$°C (dashed), with a sensor of standard deviation 0.5 K: the normal quantiles $T + 0.5\,\Phi^{-1}((i - 0.5)/12)$, $i = 1, \ldots, 12$, in an arbitrary order; red: the 24°C limit. (b) The handbook prior of the conductance, $\ln(G/6) \sim \mathcal{N}(\ln 1.25, 0.3^2)$, density $\exp[-\ln^2(G/7.5)/0.18]/(0.3\sqrt{2\pi}\,G)$; dashed: its median 7.5 kW/K; dots: the plausible values 5.556, 7.5 and 10.125 kW/K.

For three plausible values the integral is a sum. Before any reading, the three conductances of Section 3.1.2 put the hall air at $T = 18 + 30/G = 23.40$, 22.00 and 20.96°C. With the illustrative probabilities 0.25, 0.50 and 0.25 on them, the probability that the next reading crosses 24°C is

$$\Pr(y_{\text{new}} > 24°\text{C}) = 0.25 \times 0.1151 + 0.50 \times (3.2 \times 10^{-5}) + 0.25 \times (6 \times 10^{-10}) = 0.0288,$$

against $1.5 \times 10^{-5}$ at the mean conductance 7.670 kW/K, where $T = 21.91$°C: the average is 1953 times the value at the average, and almost all of it comes from the low conductance. For a Gaussian posterior $T \mid \mathbf{y} \sim \mathcal{N}(\mu, \Sigma_{TT})$ the integral has a closed form. The next reading is $y_{\text{new}} = T + \epsilon$ with the error $\epsilon \sim \mathcal{N}(0, \sigma_T^2)$ independent of $T$, and the sum of two independent normal quantities is normal with the sum of their variances, a standard result, so that

$$y_{\text{new}} \mid \mathbf{y} \sim \mathcal{N}(\mu,\ \Sigma_{TT} + \sigma_T^2), \qquad \Pr(y_{\text{new}} > c \mid \mathbf{y}) = 1 - \Phi\left(\frac{c - \mu}{\sqrt{\Sigma_{TT} + \sigma_T^2}}\right). \tag{3.3}$$

After the one reading of Chapter 1, $\mu = 23.07$°C and $\sqrt{\Sigma_{TT}} = 0.481$ K, the predictive standard deviation is $\sqrt{0.4806^2 + 0.5^2} = 0.6935$ K and the probability is $1 - \Phi(1.341) = 0.0900$, against 0.0315 for the plug-in that takes 23.07°C as the truth. In the three-value sum the average of the tail exceeds the tail at the average because $1 - \Phi((c - T)/\sigma_T)$ is convex in $T$ below the limit, where all three values lie. In the Gaussian case the reason is that the limit lies above the mean: widening the spread from 0.5 to 0.6935 K moves probability past it.

The scatter of the next reading remains even when the uncertainty about the unknown gain is removed. Table 3.1 repeats the fit of Chapter 1 with $n$ windows at 30 kW whose readings average 23.2°C, so that the state $(T, g)^{\mathsf{T}}$ carries one physics row, $n$ reading rows and the handbook prior row. The first column is the number of windows, the second and third the posterior mean and standard deviation of the hall air, the fourth the predictive standard deviation of equation (3.3), and the last two the probability that the hall air exceeds 24°C and the probability that the next reading does. The epistemic spread halves each time the windows quadruple, and the probability that the hall is above its limit falls from 0.0265 to below $10^{-10}$; the predictive spread falls only to the sensor's 0.5 K, and the probability that the next reading crosses the limit tends to $1 - \Phi(1.6) = 0.0548$, not to zero. The sensor's error is

taken as independent from window to window, with no fixed offset; an offset would set a floor under $\sqrt{\Sigma_{TT}}$ that more windows cannot remove.

The table reports the Laplace values of the state's posterior. The exact posterior of the same rows, with $T$ integrated out and a quadrature along $g$, sets the same two events beside them. The probability that the hall air exceeds 24°C is 0.0265 against the exact 0.0225 at $n = 1$, $3.76 \times 10^{-4}$ against $3.27 \times 10^{-4}$ at $n = 4$, and $4.43 \times 10^{-11}$ against $3.88 \times 10^{-11}$ at $n = 16$: the absolute gap closes while the relative gap barely moves, 14.9, 13.2 and 12.4 %. The probability that the next reading crosses is 0.0900 against 0.0813, 0.0674 against 0.0648, and 0.0583 against 0.0576. The source is the curve of the physics row: the Laplace normal is centered on the mode, 23.0701°C at $n = 1$, and the exact posterior has the mean 23.0300°C, so the exact tail above the limit is the thinner one.

Table 3.1: The hall at 30 kW after $n$ windows whose readings average 23.2°C (the fit of Chapter 1 with $n$ reading rows): the epistemic spread of the hall air shrinks, the scatter of the next reading does not.

| $n$ [–] | $\hat{T}$ [°C] | $\sqrt{\Sigma_{TT}}$ [K] | $\sqrt{\Sigma_{TT} + \sigma_T^2}$ [K] | $\Pr(T > 24°\text{C})$ [–] | $\Pr(y_{\text{new}} > 24°\text{C})$ [–] |
|---|---|---|---|---|---|
| 1 | 23.070 | 0.4806 | 0.6935 | 0.0265 | 0.0900 |
| 4 | 23.166 | 0.2476 | 0.5580 | $3.8 \times 10^{-4}$ | 0.0674 |
| 16 | 23.191 | 0.1247 | 0.5153 | $4 \times 10^{-11}$ | 0.0583 |
| 64 | 23.198 | 0.0625 | 0.5039 | $< 10^{-30}$ | 0.0557 |
| $\infty$ | 23.200 | 0 | 0.5000 | 0 | 0.0548 |

**Principle 3.1.** *Collecting more windows can shrink the uncertainty about the plant's gain without eliminating the scatter of the next reading: evidence removes epistemic uncertainty, and the aleatory part remains as long as the reading is modeled as scattering from window to window.*

### 3.1.4 Distributions of what is uncertain

Three mathematical objects represent what is uncertain, and Figure 3.2 draws one of each.

**A discrete unknown.** Suppose a cooling unit (illustrative) is either healthy or degraded, and define the indicator $F = 1$ if it is degraded and $F = 0$ if it is healthy. The assignment $\Pr(F = 1) = 0.10$, $\Pr(F = 0) = 0.90$ is a discrete probability distribution: each value receives a probability, and the probabilities sum to one. Panel (a) of the figure has the two values on the horizontal axis and their probabilities as bars.

**A continuous unknown.** Suppose the effectiveness $\eta$ of a unit's coil, the fraction of the largest possible heat transfer that it achieves, is unknown, and represent it by a density $p(\eta)$. A probability over a range is the integral of the density over that range,

$$\Pr(0.75 < \eta < 0.85) = \int_{0.75}^{0.85} p(\eta)\,\mathrm{d}\eta.$$

For a continuous variable the density itself is not a probability: it is a probability per unit of the variable, and it can exceed one. Panel (b) uses the illustrative density $\mathcal{N}(0.80, 0.04^2)$. The horizontal axis is the effectiveness and the vertical axis the density, whose peak is $1/(0.04\sqrt{2\pi}) = 9.97$; the shaded area between 0.75 and 0.85 is $\Phi(1.25) - \Phi(-1.25) = 0.789$, the probability of that range.

**A joint distribution.** Suppose an illustrative hall with a weak air path has two unknown quantities, the conductance $G$ of its air path and the supply-air temperature $T_c$, as it does when the discharge controller's own sensor is not trusted. Their *joint* distribution $p(G, T_c)$ describes the uncertainty about both together, including any dependence between them. The joint distribution is essential when an operating quantity depends on several uncertain inputs, as the hall air does,

$$T = T_c + \frac{Q}{G},$$

with the heat load $Q$ in kW and $G$ in kW/K. The individual distributions $p(G)$ and $p(T_c)$ are not always enough to determine the distribution of $T$; how $G$ and $T_c$ vary together is needed as well. Panel (c) draws two joint distributions with the same marginals, $G$ with mean 3 kW/K and standard deviation 0.15 kW/K, $T_c$ with mean 20°C and standard deviation 1 K, one with correlation $\mathrm{corr}(G, T_c) = 0$ and one with $\mathrm{corr}(G, T_c) = 0.8$. The horizontal axis is $G$, the vertical axis $T_c$, and each ellipse encloses 95 % of its distribution. Both ellipses span the same ranges of $G$ and of $T_c$, so the marginals cannot tell them apart; Section 3.6.4 shows that at $Q = 30$ kW the hall air has a standard deviation of 1.12 K under the first and 0.67 K under the second.

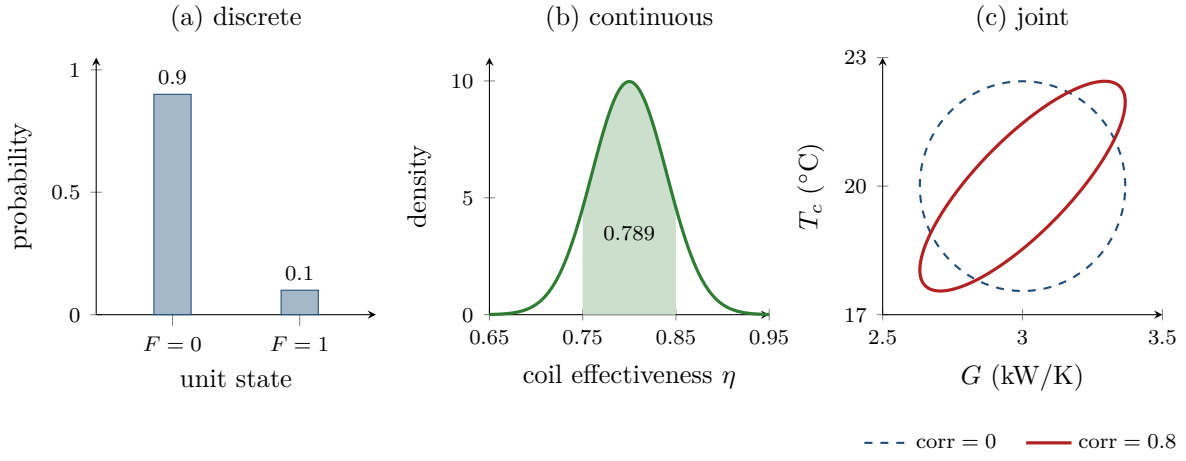



Figure 3.2: Distributions of what is uncertain (illustrative). (a) A discrete unknown: $\Pr(F = 1) = 0.10$ for a degraded cooling unit. (b) A continuous unknown: the density $\mathcal{N}(0.80, 0.04^2)$ of a coil effectiveness; shaded, $\Pr(0.75 < \eta < 0.85) = 0.789$. (c) Two joint distributions of $G$ and $T_c$ with the same normal marginals (means 3 kW/K and 20°C, standard deviations 0.15 kW/K and 1 K) and correlations 0 and 0.8; the ellipses enclose 95 % of each ($\chi^2_{2,0.95} = 5.991$).

## 3.2 Conditional probability and Bayes' theorem

Conditional probability is the foundation of every update in this book. A trip log, read run by run, shows where it comes from, why the direction of the conditioning matters, and how Bayes' theorem reverses it.

### 3.2.1 Conditional probability from a trip log

Let $H$ be the event that a loop's trip flag fires during a run and $E$ the event that the valve really stuck during the same run, established afterwards from a stroke test. The question *given that the flag fired, what is the probability that the stiction was real* is written $\Pr(E \mid H)$. The phrase "given that" restricts attention to the situations in which $H$ occurred.

Imagine 1000 comparable loop-runs of a site's trip log, gathered over its loops and its nights and recorded in Table 3.2. The rows say whether the flag fired, the columns whether real

stiction occurred, and the margins are the totals: trip and stiction together in 20 runs, the trip alone in 60, the stiction alone in 20, and neither in 900. The runs are pooled over the site's loops, not drawn from this loop's log of Chapter 2: a trip fires in 0.08 of these runs, against a probability of 0.0354 of a trip in a night for this loop, Chapter 2's value (0.0314 exact), 2.26 times as often (2.55 against the exact value).

Table 3.2: Outcomes of 1000 comparable loop-runs of a trip log (illustrative), in numbers of runs.

| | real stiction [–] | no stiction [–] | total [–] |
|---|---|---|---|
| trip fired | 20 | 60 | 80 |
| no trip | 20 | 900 | 920 |
| total | 40 | 960 | 1000 |

From these counts,

$$\Pr(H) = \frac{80}{1000} = 0.08, \qquad \Pr(E) = \frac{40}{1000} = 0.04, \qquad \Pr(H \cap E) = \frac{20}{1000} = 0.02.$$

Among the 80 runs in which the flag fired, the stiction was real in 20. Therefore

$$\Pr(E \mid H) = \frac{20}{80} = 0.25.$$

The conditional probability is 25 %, against an unconditional probability of 4 % that a run has real stiction: in these observations the trip and the stiction are strongly associated, and still three trips in four have another cause. Two flags in six weeks are therefore not two findings of stiction.

### 3.2.2 The definition

Dividing the numerator and the denominator of 20/80 by the total number of runs turns the count ratio into a ratio of probabilities,

$$\Pr(E \mid H) = \frac{20/1000}{80/1000} = \frac{\Pr(H \cap E)}{\Pr(H)}.$$

For any two events $A$ and $B$ with $\Pr(B) > 0$ the definition is

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}. \tag{3.4}$$

Conditioning on $B$ restricts the sample space to the outcomes in $B$ and renormalizes, so that the restricted probabilities again sum to one. The definition requires the conditioning event to have positive probability. Two events are *independent* when $\Pr(A \cap B) = \Pr(A)\,\Pr(B)$, which by equation (3.4) is the same as $\Pr(A \mid B) = \Pr(A)$: learning $B$ does not change the probability of $A$.

Table 3.3 works the definition for four joint probabilities with the marginals of the log held fixed, $\Pr(H) = 0.08$ and $\Pr(E) = 0.04$. The first column is the probability that a run has both a trip and real stiction, the next two are the conditional probabilities in the two directions, and the last says what kind of flag the row describes. The third row is the log of Table 3.2: $\Pr(E \mid H) = 0.02/0.08 = 0.25$ and $\Pr(H \mid E) = 0.02/0.04 = 0.50$. The second row is the independent case, in which $\Pr(E \mid H) = 0.0032/0.08 = 0.04 = \Pr(E)$: a flag that fires as often with a free valve as with a sticking one carries no information. The joint probability cannot exceed the smaller marginal, 0.04; at that bound all real stiction raises the flag.

Table 3.3: Conditional probabilities from equation (3.4) for four joint probabilities of a trip and real stiction, with $\Pr(H) = 0.08$ and $\Pr(E) = 0.04$ held fixed.

| $\Pr(H \cap E)$ [–] | $\Pr(E \mid H)$ [–] | $\Pr(H \mid E)$ [–] | Kind of flag |
|---|---|---|---|
| 0 | 0 | 0 | never fires with real stiction |
| 0.0032 | 0.04 | 0.08 | independent of the stiction: no information |
| 0.02 | 0.25 | 0.50 | the logged flag: positively associated |
| 0.04 | 0.50 | 1.00 | all real stiction raises the flag |

### 3.2.3 The multiplication rule

Multiplying equation (3.4), written for $\Pr(B \mid A)$, by $\Pr(A)$ gives the multiplication rule,

$$\Pr(A \cap B) = \Pr(B \mid A)\ \Pr(A). \tag{3.5}$$

Conditioning in the other direction gives $\Pr(A \cap B) = \Pr(A \mid B)\ \Pr(B)$, and the two expressions for the same intersection are equal:

$$\Pr(B \mid A)\ \Pr(A) = \Pr(A \mid B)\ \Pr(B). \tag{3.6}$$

For the log, $0.25 \times 0.08 = 0.02 = 0.50 \times 0.04$. This relation is the starting point of Bayes' theorem.

**Reversing the conditioning.** The two conditional probabilities of the log differ: $\Pr(E \mid H) = 0.25$, but $\Pr(H \mid E) = 20/40 = 0.50$. The first asks about the stiction among the runs in which the flag fired, the fraction of trips that are real stiction; the second asks about the flag among the runs with real stiction, the flag's probability of detection per run. Half the real stiction is missed because the stick is small enough to stay inside the flag's error band, and the on-delay ignores the short excursions a small stick causes. Equation (3.6) shows that they are equal only when $\Pr(H) = \Pr(E)$.

**Common mistake**

Treating $\Pr(A \mid B)$ and $\Pr(B \mid A)$ as the same number. Reversing the conditioning changes the question. In fault diagnosis the probability of a trip given a fault is known from the design of the flag, and the probability of a fault given a trip is what the engineer needs; the second follows from the first only through Bayes' theorem and the prior probability of the fault.

### 3.2.4 Bayes' theorem

Dividing both sides of equation (3.6) by $\Pr(B) > 0$ gives Bayes' theorem for discrete events,

$$\Pr(A \mid B) = \frac{\Pr(B \mid A)\ \Pr(A)}{\Pr(B)}. \tag{3.7}$$

For the log, the flag's probability of detection and the base rates give the fraction of trips that are real stiction, $\Pr(E \mid H) = 0.50 \times 0.04/0.08 = 0.25$, the value read from the counts. In words,

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}},$$

and the four terms have distinct roles when $A$ is a hypothesis about the plant and $B$ an observation. The *prior* $\Pr(A)$ is what was believed about the hypothesis before the observation. The *likelihood* $\Pr(B \mid A)$ is how probable the observation is when the hypothesis holds. The *evidence* $\Pr(B)$ is the overall probability of the observation, averaged over the hypotheses; it is the normalizing constant of the posterior and, for a whole model $M$, the quantity the notation writes $p(\mathbf{y} \mid M)$. The *posterior* $\Pr(A \mid B)$ is what is believed about the hypothesis after the observation is accounted for. The engineer's question "this flag just fired; what is the probability it is real" is this computation, and Chapter 26 answers it on the running loop.

## 3.3 Bayesian diagnosis of a cooling fault

Bayes' theorem turns the known behavior of a trip flag into the probability of the fault the flag is meant to detect, and the answer is far below the flag's probability of detection; a silent flag is evidence too.

### 3.3.1 The diagnosis problem

The hall's cooling unit also carries a valve-position trip, which watches the valve opening against what the load calls for and is designed to detect degradation of the unit. Define the events

$$F = \{\text{the cooling unit is degraded}\}, \qquad H = \{\text{the valve-position trip occurs}\}.$$

The maintenance log suggests that 5 % of runs have a degraded unit, $\Pr(F) = 0.05$. The trip activates in 90 % of degraded conditions, $\Pr(H \mid F) = 0.90$, its probability of detection; it also fires falsely, in 10 % of healthy runs, $\Pr(H \mid \neg F) = 0.10$, its false-trip probability. A data sheet that states the first number invites the reading that a trip means a degraded unit with probability 0.90. A trip is observed, and the question is the probability that the unit is in fact degraded, $\Pr(F \mid H)$.

The first step writes Bayes' theorem, equation (3.7), for the hypothesis $F$ and the observation $H$:

$$\Pr(F \mid H) = \frac{\Pr(H \mid F)\,\Pr(F)}{\Pr(H)}.$$

The numerator is known, $\Pr(H \mid F)\,\Pr(F) = 0.90 \times 0.05 = 0.045$. The denominator, the probability of a trip, is not yet known.

### 3.3.2 The total probability of a trip

A trip can occur in two mutually exclusive situations: the unit is degraded and the trip activates, or the unit is healthy and the trip activates falsely. In this simplified model the two exhaust the cases, so the trip event is the union of two disjoint events and its probability is their sum,

$$\Pr(H) = \Pr(H \cap F) + \Pr(H \cap \neg F).$$

The multiplication rule, equation (3.5), writes each intersection as a likelihood times a prior:

$$\begin{aligned} \Pr(H) &= \Pr(H \mid F)\,\Pr(F) + \Pr(H \mid \neg F)\,\Pr(\neg F) \\ &= 0.90 \times 0.05 + 0.10 \times 0.95 = 0.045 + 0.095 = 0.14. \end{aligned} \tag{3.8}$$

Trips occur in 14 % of runs; two thirds of them, 0.095 of the 0.14, come from runs with a healthy unit.

### 3.3.3 The posterior fault probability

The third step divides,

$$\Pr(F \mid H) = \frac{0.045}{0.14} = 0.3214. \tag{3.9}$$

The conclusion does not rest on a probability of 0.90: at 0.32 a healthy unit is the likelier explanation, and whether a technician is worth sending depends on what a visit costs against what a missed degradation costs. Figure 3.3 compares the probability of degradation before and after the observation. Each row is one state of knowledge, the horizontal scale is a probability from 0 to 1, and the green part of each bar is the probability of degradation. The trip raises it from 5 % to 32.1 %; it increases suspicion substantially, but it does not establish that a fault is present. The third row is the case without a trip, derived in Section 3.3.4.

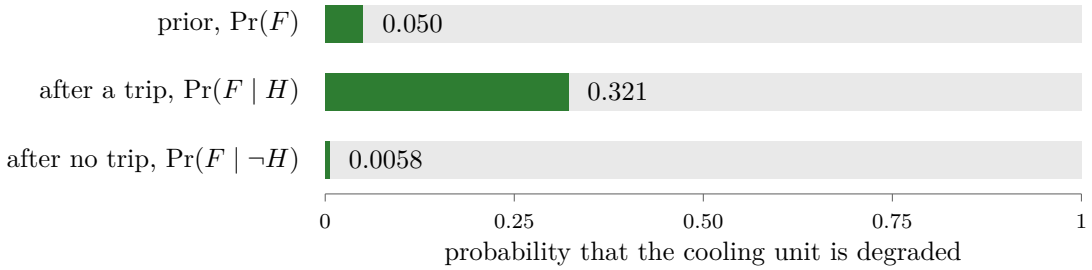



Figure 3.3: What a trip tells about degradation. Prior $\Pr(F) = 0.05$ per run; trip with probability of detection $\Pr(H \mid F) = 0.90$ and false-trip probability $\Pr(H \mid \neg F) = 0.10$. After a trip, $0.045/0.14 = 0.321$ (equation (3.9)); after no trip, $0.005/0.86 = 0.0058$ (equation (3.11)).

**The base rate.** The answer is not 90 %. The trip detects 90 % of degraded conditions, but degraded conditions are uncommon, and many trips occur while the unit is healthy. Table 3.4 makes this concrete for 1000 runs. The rows are the actual condition of the unit, the columns whether the trip fired, and each entry is the expected number of runs: of the 50 degraded runs, 90 %, or 45, raise the trip; of the 950 healthy runs, 10 %, or 95, raise it falsely. Among the 140 runs with a trip only 45 have a degraded unit, so $\Pr(F \mid H) = 45/140 = 0.3214$. The table and Bayes' theorem give the same result, because the table is equation (3.8) multiplied by 1000.

Table 3.4: Expected outcomes of 1000 runs with $\Pr(F) = 0.05$, $\Pr(H \mid F) = 0.90$ and $\Pr(H \mid \neg F) = 0.10$, in numbers of runs.

| Actual condition | trip [–] | no trip [–] | total [–] |
|---|---|---|---|
| degraded | 45 | 5 | 50 |
| healthy | 95 | 855 | 950 |
| total | 140 | 860 | 1000 |

**The odds form.** Writing Bayes' theorem for $F$ and for $\neg F$ with the same observation and dividing the two, the evidence $\Pr(H)$ cancels:

$$\frac{\Pr(F \mid H)}{\Pr(\neg F \mid H)} = \frac{\Pr(H \mid F)}{\Pr(H \mid \neg F)} \cdot \frac{\Pr(F)}{\Pr(\neg F)}. \tag{3.10}$$

Posterior odds equal the likelihood ratio times the prior odds. Here the likelihood ratio is $0.90/0.10 = 9$ and the prior odds are $0.05/0.95 = 0.05263$, so the posterior odds are 0.4737 and

the posterior probability is $0.4737/(1+0.4737) = 0.3214$, as before. The trip multiplies the odds of degradation by nine whatever the base rate; the base rate decides where that multiplication starts.

### 3.3.4 The case without a trip

The same derivation applies to the observation that the trip does not activate. Bayes' theorem gives

$$\Pr(F \mid \neg H) = \frac{\Pr(\neg H \mid F)\ \Pr(F)}{\Pr(\neg H)},$$

with $\Pr(\neg H \mid F) = 1 - 0.90 = 0.10$, the probability that a degraded unit goes undetected, and, by the same total probability as equation (3.8),

$$\Pr(\neg H) = 0.10 \times 0.05 + 0.90 \times 0.95 = 0.005 + 0.855 = 0.86.$$

Therefore

$$\Pr(F \mid \neg H) = \frac{0.005}{0.86} = 0.00581. \tag{3.11}$$

The absence of a trip reduces the fault probability from 5 % to about 0.58 %, the third row of Figure 3.3; in odds, the silence multiplies the prior odds by $0.10/0.90 = 1/9$. In this example a trip substantially raises suspicion, and its absence is considerable evidence that the unit is healthy.

**Interpretation.** Bayesian diagnosis does not only ask whether a reading crossed a threshold. It evaluates how strongly the observation supports competing explanations, weighted by their prior plausibility. The distinction is central on a running loop because many abnormal readings are not uniquely associated with one physical fault: a high hall temperature can be caused by a reduced cooling capacity, an increased load, a warmer supply air or a biased sensor. Each of these is a hypothesis with its own prior and its own likelihood for the same reading, and the inference becomes a comparison among them, the diagnosis category of Chapter 2, which Chapters 26 and 27 develop as fault detection and isolation.

## 3.4 Marginalization and dependence

The probability of a trip in Section 3.3.2 was computed without knowing whether the unit was degraded. The same operation, for discrete and continuous unknowns and for nuisance quantities, is what lets a shared pump make two loops dependent, and what an independence assumption costs.

### 3.4.1 Discrete marginalization

Suppose a cooling unit (illustrative) can be in one of three operating conditions, healthy, degraded or failed, and let $S$ denote the condition present. Table 3.5 gives, for each state, its probability and the probability that the trip $H$ occurs in it; the last two columns are computed below.

A trip can arise under any of the three states. Because the states are mutually exclusive and exhaustive, the events $H \cap \{S = s\}$ are disjoint and their union is $H$, so $\Pr(H) = \sum_s \Pr(H \cap S = s)$. The multiplication rule writes each term as a likelihood times a prior:

$$\Pr(H) = \sum_s \Pr(H \mid S = s)\ \Pr(S = s). \tag{3.12}$$

Table 3.5: Three operating conditions of a cooling unit: probabilities, trip likelihoods, the joint probabilities of state and trip, and the probabilities of the states given a trip.

| State $s$ | $\Pr(S=s)$ [–] | $\Pr(H \mid S=s)$ [–] | $\Pr(H \cap S=s)$ [–] | $\Pr(S=s \mid H)$ [–] |
|---|---|---|---|---|
| healthy | 0.85 | 0.05 | 0.0425 | 0.272 |
| degraded | 0.12 | 0.70 | 0.0840 | 0.538 |
| failed | 0.03 | 0.99 | 0.0297 | 0.190 |
| total | 1.00 | | 0.1562 | 1.000 |

This is the *law of total probability*. Substituting,

$$\Pr(H) = 0.05 \times 0.85 + 0.70 \times 0.12 + 0.99 \times 0.03 = 0.0425 + 0.0840 + 0.0297 = 0.1562.$$

The probability of a trip has been computed without knowing which state is present. That is what marginalization does: it accounts for the uncertainty in variables that are not part of the final question. The same sum is the denominator of Bayes' theorem over the three states, and dividing each joint probability by it gives the last column of the table: given a trip, the unit is degraded with probability $0.0840/0.1562 = 0.538$, healthy with 0.272 and failed with 0.190.

### 3.4.2 Continuous marginalization

When the unknown is continuous, the sum becomes an integral. Let the rate $\lambda$ at which the loop's trip flag fires be uncertain and let $\mathbf{y}$ be the observed evidence; their joint distribution is $p(\lambda, \mathbf{y}) = p(\mathbf{y} \mid \lambda)\, p(\lambda)$. The probability density of the observations is obtained by integrating over every possible value of the rate,

$$p(\mathbf{y}) = \int p(\mathbf{y} \mid \lambda)\, p(\lambda)\, \mathrm{d}\lambda, \tag{3.13}$$

the *marginal likelihood* of the observations. Bayes' theorem for a continuous unknown then reads

$$p(\lambda \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \lambda)\, p(\lambda)}{\int p(\mathbf{y} \mid \lambda')\, p(\lambda')\, \mathrm{d}\lambda'}, \tag{3.14}$$

where $\lambda'$ is a dummy variable of integration. The numerator evaluates the relative plausibility of one particular rate; the denominator ensures that the posterior integrates to one.

For the loop's trip log of Chapter 2, $k = 2$ trips in $T_{\mathrm{tot}} = 1000\,\mathrm{h}$ of logged operation under the Poisson model $\Pr(k \mid \lambda) = (\lambda T_{\mathrm{tot}})^k \mathrm{e}^{-\lambda T_{\mathrm{tot}}}/k!$ and the prior $\mathrm{Gamma}(a, b) = \mathrm{Gamma}(2, 500\,\mathrm{h})$, with density $b^a \lambda^{a-1} \mathrm{e}^{-b\lambda}/\Gamma(a)$, the standard gamma integral $\int_0^\infty \lambda^{a-1} \mathrm{e}^{-s\lambda}\, \mathrm{d}\lambda = \Gamma(a)/s^a$ with $s = b + T_{\mathrm{tot}}$ evaluates equation (3.13) in closed form:

$$\Pr(k \mid T_{\mathrm{tot}}) = \frac{T_{\mathrm{tot}}^k}{k!} \frac{b^a}{\Gamma(a)} \frac{\Gamma(a+k)}{(b+T_{\mathrm{tot}})^{a+k}} = \frac{1000^2}{2} \cdot \frac{500^2}{1} \cdot \frac{3!}{1500^4} = 0.1481.$$

Before the log, the model gave a probability of 0.148 to exactly two trips in 1000 h; dividing the numerator of equation (3.14) by it gives the posterior $\mathrm{Gamma}(4, 1500\,\mathrm{h})$. The probability of a clean night run, no trip in the next $\tau = 12\,\mathrm{h}$, is the same operation applied to the event of no trip: its conditional probability $\mathrm{e}^{-\lambda\tau}$ averaged over the posterior of the rate. Combining the two exponentials and using the same gamma integral with $s = b + \tau$, for the posterior $\mathrm{Gamma}(a, b)$,

$$\Pr(\text{no trip in } \tau \mid \mathbf{y}) = \frac{b^a}{\Gamma(a)} \int_0^\infty \lambda^{a-1}\, \mathrm{e}^{-(b+\tau)\lambda}\, \mathrm{d}\lambda = \frac{b^a}{\Gamma(a)} \frac{\Gamma(a)}{(b+\tau)^a} = \left(\frac{b}{b+\tau}\right)^a, \tag{3.15}$$

which for Gamma(4, 1500 h) is $(1500/1512)^4 = 0.96863$. Chapter 2 reports this probability from the posterior of the state, the same average over the normal of the log-rate by the five-node sum of equation (2.41): 0.96457, with the gamma integral 0.96863 as its exact value beside it, a gap of 0.0041. A trip in the night is then 0.0354 (0.0314 exact). The plug-in at the posterior mean rate, $\mathrm{e}^{-12\times4/1500} = 0.9685$, sits 0.0001 from the exact value and answers a different question, a trip whose rate is known. The predictive probability of equation (3.2) is the same operation applied to the next reading: the conditional probability averaged over the posterior of the hall air.

### 3.4.3 Nuisance quantities

Suppose the log-conductance $g$ of the hall is being estimated from the hall reading and the load meter, but the model also contains the uncertain calibration offset $b_Q$ of the load meter, in kW. The full posterior is $p(g, b_Q \mid \mathbf{y})$, and the engineer wants the distribution of $g$ alone. It is obtained by integrating $b_Q$ out,

$$p(g \mid \mathbf{y}) = \int p(g, b_Q \mid \mathbf{y}) \, \mathrm{d}b_Q. \tag{3.16}$$

The offset does not have to be treated as known; its uncertainty enters the answer through the integration. This matters most in physics-constrained inference, where many unknown state entries influence the quantity studied: the marginal of one entry $z_i$ is the posterior of the whole state $\mathbf{z}$ integrated over all its other entries.

**Marginalization against optimization.** These two operations are different and must not be confused. Marginalization, equation (3.16), accumulates probability across all values of $b_Q$. Optimization over the nuisance quantity, $\max_{b_Q} p(g, b_Q \mid \mathbf{y})$, selects the best-fitting offset for each conductance. They generally give different results. A discrete illustration makes the difference plain. Let the conductance be either low or high and the offset take one of three values, with joint posterior probabilities 0.30, 0.05, 0.05 for a low conductance and 0.20, 0.20, 0.20 for a high one (illustrative). Marginalization gives Pr(low) = 0.40 and Pr(high) = 0.60; maximization over $b_Q$ gives 0.30 against 0.20 and prefers the low conductance. The low conductance owes its single peak to one value of the offset; the high one is supported by every value. The distinction returns when later chapters compare a maximum a posteriori estimate with the marginals of the posterior.

### 3.4.4 Two loops on one chilled-water pump

Many tuning and trip calculations assume that the loops of a site behave independently. Let $F_A$ and $F_B$ be the events that loop A and loop B trip in the same run. If each has probability 0.05, independence gives $\Pr(F_A \cap F_B) = 0.05^2 = 0.0025$ for the probability that both trip at once. The conclusion is invalid if the loops share an important disturbance. Consider two loops whose valves draw chilled water through one pump. Their supply airs are taken as equal, and a sag the valves can absorb lifts them only briefly, while the integral catches up; it holds them up once the valves reach their stops. Let $C$ be the event that the pump's flow sags during a run, through a pump trip or a staging delay, with $\Pr(C) = 0.05$. The event $C$ counts every run in which the flow sags enough to disturb both loops, staging delays and partial losses as well as trips, so it is more frequent than the pump trips $q_S$ of the two-loops example of Chapter 2; and it raises each loop's trip probability, to 0.60 and 0.50 below, rather than forcing both. Figure 3.4 draws the dependency: the box at the top is the shared disturbance, the two boxes below are the trip events of the loops, and each arrow says that the probability of the event at its head depends on the condition at its tail. There is no arrow between the loops.

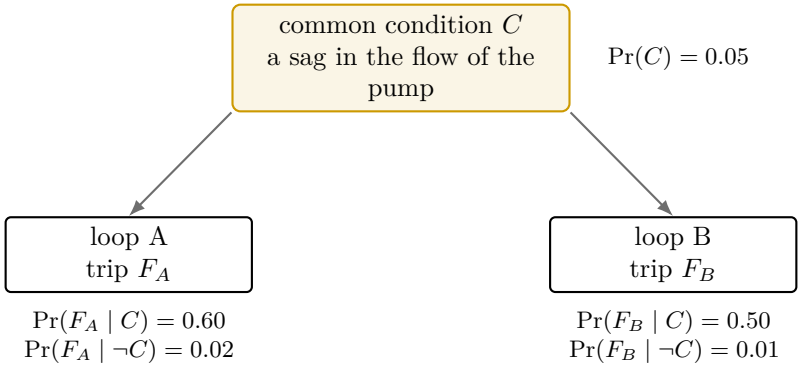



Figure 3.4: A probabilistic dependency graph for two loops on one chilled-water pump. The shared condition $C$, a sag in the pump's flow, influences both trip events; the absence of an arrow between the loops states that the trips are conditionally independent given $C$. The probabilities are those of Table 3.6.

The trip probabilities of the loops depend on whether the pump's flow sags. Table 3.6 lists them: the rows are the two conditions and the columns the probability that each loop trips in a run under that condition.

Table 3.6: Probabilities that each loop trips in a run, with and without a sag in the flow of the shared pump (illustrative); $\Pr(C) = 0.05$.

| Condition | $\Pr(F_A \mid \cdot)$ [–] | $\Pr(F_B \mid \cdot)$ [–] |
|---|---|---|
| flow sags, $C$ | 0.60 | 0.50 |
| normal flow, $\neg C$ | 0.02 | 0.01 |

Assume that the two trips are *conditionally independent given* $C$:

$$\Pr(F_A \cap F_B \mid C) = \Pr(F_A \mid C)\,\Pr(F_B \mid C), \qquad \Pr(F_A \cap F_B \mid \neg C) = \Pr(F_A \mid \neg C)\,\Pr(F_B \mid \neg C).$$

Once the condition of the pump is known, a trip in one loop says nothing more about the other. This does not mean that the trips are unconditionally independent, as the next subsection shows.

### 3.4.5 The joint trip probability and the independence error

The law of total probability, equation (3.12), applied to the event $F_A \cap F_B$ with the two conditions $C$ and $\neg C$ as the states, gives

$$\Pr(F_A \cap F_B) = \Pr(F_A \cap F_B \mid C)\,\Pr(C) + \Pr(F_A \cap F_B \mid \neg C)\,\Pr(\neg C).$$

Under conditional independence each conditional joint probability is a product,

$$\Pr(F_A \cap F_B) = 0.60 \times 0.50 \times 0.05 + 0.02 \times 0.01 \times 0.95 = 0.015 + 0.00019,$$

so that

$$\Pr(F_A \cap F_B) = 0.01519. \tag{3.17}$$

The probability that both loops trip in the same run is about 1.52 %, and almost all of it, 0.015 of 0.01519, comes from the pump.

The marginal trip probabilities follow from the same law,

$$\Pr(F_A) = 0.60 \times 0.05 + 0.02 \times 0.95 = 0.049, \qquad \Pr(F_B) = 0.50 \times 0.05 + 0.01 \times 0.95 = 0.0345.$$

If the trips are incorrectly treated as independent, the joint probability is the product of the marginals,

$$\Pr(F_A \cap F_B) \overset{\text{incorrect}}{=} 0.049 \times 0.0345 = 0.00169,$$

or 0.169 %. The correct value, 1.519 %, is 8.99 times larger: ignoring the shared pump underestimates the probability that both loops trip at once by a factor of about nine in this illustrative site, and the independent product is only 11 % of the correct value. With marginals of the same size, 0.049 and 0.0345 here against 0.05 and 0.03 there, this factor of nine sits beside the factor of 7.2 that Chapter 2 finds for two loops both tripping at a pump-trip probability $q_S = 0.01$.

**Common mistake**

Multiplying the trip probabilities of two loops to obtain the probability that both trip at once. The product is the joint probability only when the trips are independent; a shared pump, a shared outdoor condition or a shared control sequence makes them dependent, and the night on which both loops need the engineer at the same time is much more likely than the product suggests.

**Principle 3.2.** *Understanding what the loops of a plant share can matter more than estimating the trip probability of each one with greater precision.*

### 3.4.6 Conditional independence given the pump

Two random variables $X_1$ and $X_2$ are conditionally independent given $C$ when, for every value $c$ of $C$,

$$p(x_1, x_2 \mid c) = p(x_1 \mid c)\, p(x_2 \mid c), \tag{3.18}$$

written $X_1 \perp X_2 \mid C$. Dividing both sides by $p(x_1 \mid c)$ and using the definition of conditional probability on the left gives the equivalent statement $p(x_2 \mid x_1, c) = p(x_2 \mid c)$: after conditioning on $C$, learning $X_1$ gives no additional information about $X_2$ under the assumed model. The two loops in physical form are the case: with the two supply airs taken as equal, at $T_c$, the hall air temperatures $T_A = T_c + Q_A/G_A$ and $T_B = T_c + Q_B/G_B$ depend on separate loads and conductances, so that given $T_c$ they are independent; with $T_c$ unknown, a reading in the hall of loop A informs $T_c$ and through it the hall of loop B.

When $C$ is unknown, observing $X_1$ informs $C$, and that changes what is believed about $X_2$. If loop A trips, the pump becomes suspect: by Bayes' theorem,

$$\Pr(C \mid F_A) = \frac{\Pr(F_A \mid C)\ \Pr(C)}{\Pr(F_A)} = \frac{0.60 \times 0.05}{0.049} = 0.612,$$

up from 0.05. The probability that loop B also trips is then the law of total probability conditioned on $F_A$, with conditional independence used to drop $F_A$ from $\Pr(F_B \mid C, F_A)$:

$$\begin{aligned}\Pr(F_B \mid F_A) &= \Pr(F_B \mid C)\ \Pr(C \mid F_A) + \Pr(F_B \mid \neg C)\ \Pr(\neg C \mid F_A) \\ &= 0.50 \times 0.612 + 0.01 \times 0.388 = 0.310,\end{aligned}$$

which equals $\Pr(F_A \cap F_B)/\Pr(F_A) = 0.01519/0.049$, against the unconditional $\Pr(F_B) = 0.0345$. A trip in one loop raises the trip probability of the other ninefold, through the pump they share: the engineer who sees a trip in loop A has reason to look at loop B, and to look at the pump before the valve of loop A. Conditional independence is therefore not unconditional independence: the variables are independent only once the common cause is known.

## 3.5 Conditional independence and a continuous update in a physical model

The same rules apply when the unknowns are physical quantities tied by a conservation relation. On the equipment hall they give the posterior of a state vector under Gaussian rows, the probability that the hall exceeds its limit after one reading, and, in the two-entry hall of Chapter 1, a reading that reaches the conductance through the physics row.

### 3.5.1 The hall as a probability model

The hall of the earlier chapters satisfies the steady heat balance

$$T = T_c + \frac{Q}{G}, \tag{3.19}$$

with the hall air $T$ and the supply air $T_c$ in °C, the heat load $Q$ in kW and the conductance $G$ of the air path in kW/K. A sensor reads the hall air with an additive error,

$$y_T = T + \epsilon, \qquad \epsilon \sim \mathcal{N}(0, \sigma_T^2),$$

where $\epsilon$ is the measurement error and $\sigma_T$ its standard deviation in K. Figure 3.5 draws the system twice. Panel (a) is the plant graph: the source supplies the load $Q$ to the hall node, which carries the free entry $T$; the green path to the supply-air node carries the conductance, held in the state as the log-conductance $g = \ln(G/G_{\text{ref}})$; the supply-air node is pinned at the set point $T_c$ by the unit's discharge controller; and the dashed arrow is the reading $y_T$. Panel (b) is the dependency graph of the probability model of this section, in which $Q$, $G$ and $T_c$ are all treated as uncertain: the load, the conductance and the supply air determine the modeled hall air, and the reading depends on that temperature alone.

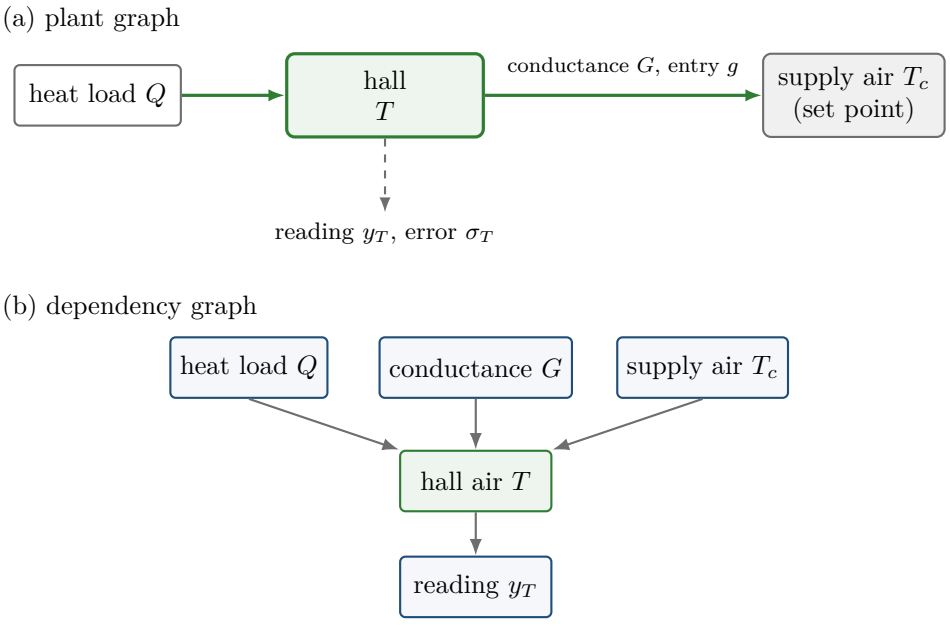



Figure 3.5: The hall in two drawings. (a) The plant graph: free entry $T$ at the hall node and $g = \ln(G/G_{\text{ref}})$ on the air path; supply air pinned at the set point $T_c$ by the discharge controller; source $Q$; reading $y_T$ observed (dashed). (b) The dependency graph of the factorized model of equation (3.21), in which $Q$, $G$ and $T_c$ carry priors: arrows point from a quantity to what depends on it, and the reading depends on the inputs only through $T$.

If the measurement error is independent of the physical inputs, the reading is conditionally independent of those inputs once the true temperature is known:

$$p(y_T \mid T, G, Q, T_c) = p(y_T \mid T). \tag{3.20}$$

This does not make the reading uninformative about the conductance. The reading informs $T$, and the physical relation (3.19) connects $T$ to $G$: evidence propagates from the sensor reading to the temperature and from the temperature to the conductance. An observation of one quantity updates the beliefs about every quantity connected to it through shared physical constraints, and this is one of the main reasons why inference over a graph is useful.

**The factorized probability model.** Suppose $G$, $Q$ and $T_c$ are uncertain inputs with a joint prior $p(G, Q, T_c)$, which relaxes the pins of the plant graph and carries $Q$ and $T_c$ as free entries. The chain rule, the multiplication rule applied twice, and equation (3.20) give the joint model

$$p(G, Q, T_c, T, y_T) = p(G, Q, T_c)\, p(T \mid G, Q, T_c)\, p(y_T \mid T). \tag{3.21}$$

The physics factor is written as a Gaussian around the heat balance, $p(T \mid G, Q, T_c) = \mathcal{N}(T;\ T_c + Q/G,\ \sigma_p^2)$, with a model tolerance $\sigma_p$ in K: a physical discrepancy is a finite $\sigma_p$, and the exact deterministic balance of equation (3.19) is the limit $\sigma_p \to 0$, in which the factor becomes the Dirac delta $\delta(T - T_c - Q/G)$. The joint distribution is the complete probability model, and the posterior wanted is $p(G, Q, T_c, T \mid y_T)$, over the whole state with $T$ kept in it, which answers questions about the conductance, the temperature, the operating margins and every other quantity of the model. Each of its three factors is a row over that state: the prior, the physics row with tolerance $\sigma_p$ and the reading row; the derivations of this section keep all of them as rows and never remove an entry by substitution. In the limit $\sigma_p \to 0$, integrating $T$ out against the delta gives the same posterior of the inputs as $p(G, Q, T_c \mid y_T) \propto p(G, Q, T_c)\, p\big(y_T \mid T = T_c + Q/G\big)$: every combination of inputs is weighted by how well the temperature it implies explains the reading.

### 3.5.2 Prior and reading for the hall air

The simplest continuous case has one unknown, the true hall air temperature $T$, so that the state is this one number,

$$T \in \mathbb{R}.$$

The general forms of the book write the state as a vector $\mathbf{z}$; here it has the single entry $T$, so every quantity of the fit below is a number: each row's residual and its slope in $T$, the objective's slope and curvature, the information and the posterior variance. On an illustrative hall running warmer than the series' hall, a commissioning model gives an uncertain prior estimate,

$$T \sim \mathcal{N}(\mu_0, \sigma_0^2) = \mathcal{N}(26, 2^2), \qquad \mu_0 = 26^\circ\mathrm{C}, \quad \sigma_0 = 2\,\mathrm{K}.$$

A sensor reports $y_T = 27^\circ$C, and its error is Gaussian with $\sigma_T = 0.5\,\mathrm{K}$, so that $y_T \mid T \sim \mathcal{N}(T, 0.5^2)$. The question is the posterior distribution of the true temperature after the reading. Bayes' theorem for a continuous unknown, with the evidence dropped because it does not depend on the state, gives

$$p(T \mid y_T) \propto p(y_T \mid T)\, p(T),$$

with the Gaussian prior and the Gaussian likelihood

$$p(T) \propto \exp\left[-\frac{(T - \mu_0)^2}{2\sigma_0^2}\right], \qquad p(y_T \mid T) \propto \exp\left[-\frac{(y_T - T)^2}{2\sigma_T^2}\right].$$

Multiplying,

$$p(T \mid y_T) \propto \exp\left[-\frac{1}{2}\left(\frac{(T-\mu_0)^2}{\sigma_0^2}+\frac{(y_T-T)^2}{\sigma_T^2}\right)\right]. \tag{3.22}$$

Each term of the exponent is the square of a difference divided by its standard deviation. The two differences are the residuals of the problem, one prior row and one reading row, each already divided by its standard deviation,

$$r_1(T)=\frac{T-\mu_0}{\sigma_0}, \qquad r_2(T)=\frac{T-y_T}{\sigma_T}, \tag{3.23}$$

and equation (3.22) reads $p(T \mid y_T) \propto \exp[-\frac{1}{2}(r_1^2+r_2^2)] = \mathrm{e}^{-\mathcal{L}(T)}$, with the objective $\mathcal{L}(T) = \frac{1}{2}(r_1^2+r_2^2)$ of the notation chapter. Both rows are linear in $T$, each with a constant slope $j_k = \mathrm{d}r_k/\mathrm{d}T$: $j_1 = 1/\sigma_0$ and $j_2 = 1/\sigma_T$.

**Gaussian rows over any state.** The same step holds for a state of any size and any number of rows. The joint model factorizes by the chain rule, as equation (3.21) does for the hall: a prior factor over the inputs, a physics factor for each physical relation, and a reading factor for each reading, which by conditional independence, equation (3.20), depends on the state only through the quantity read. Row $k$ is one such factor written as a residual: the difference between the quantity computed from the state and what was observed, or what the prior or the physics expects, divided by its standard deviation $\sigma_k$. A Gaussian factor is $\exp[-\frac{1}{2}r_k(\mathbf{z})^2]$ times a normalizer, $1/(\sqrt{2\pi}\,\sigma_k)$, that is free of the state. Multiplying the factors and dropping the normalizers and the evidence, Bayes' theorem gives

$$p(\mathbf{z} \mid \mathbf{y}) \propto \prod_k \exp\left[-\tfrac{1}{2}r_k(\mathbf{z})^2\right] = \exp\left[-\tfrac{1}{2}r(\mathbf{z})^{\mathsf{T}}r(\mathbf{z})\right] = \mathrm{e}^{-\mathcal{L}(\mathbf{z})}, \tag{3.24}$$

where $r(\mathbf{z})$ stacks the reading rows, the prior rows and the physics rows with their tolerances $\sigma_p$. For the one-entry hall the product has the two factors of equation (3.22) and the sum of squares is the number $r_1^2+r_2^2$; the vector is needed when the state has several entries, as in Section 3.5.6 and in the hall of Chapter 4. A Gaussian prior over several entries, $\mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ with correlated entries, enters as the rows $\mathbf{L}_0(\mathbf{z}-\boldsymbol{\mu}_0)$ with $\mathbf{L}_0^{\mathsf{T}}\mathbf{L}_0 = \boldsymbol{\Sigma}_0^{-1}$, since then $(\mathbf{z}-\boldsymbol{\mu}_0)^{\mathsf{T}}\boldsymbol{\Sigma}_0^{-1}(\mathbf{z}-\boldsymbol{\mu}_0)$ is the sum of their squares.

### 3.5.3 Completion of the square in the state

Suppose the rows of a state with several entries are linear in it, as both rows of equation (3.23) are in $T$. Then, for any point $\mathbf{z}_0$,

$$r(\mathbf{z}) = r(\mathbf{z}_0) + \mathbf{J}\,(\mathbf{z}-\mathbf{z}_0)$$

holds exactly, with the Jacobian $\mathbf{J} = \partial r/\partial \mathbf{z}$, a matrix of $m$ rows and $n$ columns for $m$ rows and $n$ entries. Write $\boldsymbol{\Delta} = \mathbf{z}-\mathbf{z}_0$ and $r_0 = r(\mathbf{z}_0)$. The objective expands as

$$\mathcal{L}(\mathbf{z}) = \tfrac{1}{2}(r_0+\mathbf{J}\boldsymbol{\Delta})^{\mathsf{T}}(r_0+\mathbf{J}\boldsymbol{\Delta}) = \tfrac{1}{2}r_0^{\mathsf{T}}r_0 + \boldsymbol{\Delta}^{\mathsf{T}}\mathbf{J}^{\mathsf{T}}r_0 + \tfrac{1}{2}\boldsymbol{\Delta}^{\mathsf{T}}\mathbf{J}^{\mathsf{T}}\mathbf{J}\,\boldsymbol{\Delta}.$$

Its gradient, $\mathbf{J}^{\mathsf{T}}r_0 + \mathbf{J}^{\mathsf{T}}\mathbf{J}\,\boldsymbol{\Delta}$, vanishes at the step $\hat{\boldsymbol{\Delta}}$ that solves the *normal equations*,

$$(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\hat{\boldsymbol{\Delta}} = -\mathbf{J}^{\mathsf{T}}r(\mathbf{z}_0), \qquad \hat{\mathbf{z}} = \mathbf{z}_0 + \hat{\boldsymbol{\Delta}}. \tag{3.25}$$

Writing $\mathbf{J}^{\mathsf{T}}r_0 = -\mathbf{J}^{\mathsf{T}}\mathbf{J}\,\hat{\boldsymbol{\Delta}}$ in the expansion completes the square,

$$\mathcal{L}(\mathbf{z}) = \tfrac{1}{2}(\boldsymbol{\Delta}-\hat{\boldsymbol{\Delta}})^{\mathsf{T}}\mathbf{J}^{\mathsf{T}}\mathbf{J}\,(\boldsymbol{\Delta}-\hat{\boldsymbol{\Delta}}) + \text{constant} = \tfrac{1}{2}(\mathbf{z}-\hat{\mathbf{z}})^{\mathsf{T}}\mathbf{J}^{\mathsf{T}}\mathbf{J}\,(\mathbf{z}-\hat{\mathbf{z}}) + \text{constant},$$

where the constant collects the terms free of $\mathbf{z}$. By equation (3.24) the posterior is the exponential of minus this quadratic form, the density of a multivariate Gaussian with mean $\hat{\mathbf{z}}$ and covariance

$$\boldsymbol{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}. \tag{3.26}$$

The matrix $\mathbf{J}^\mathsf{T}\mathbf{J}$ is the information matrix $\boldsymbol{\Lambda}$ of the normalized rows. When the rows are written unnormalized, each in its own unit with the weight $w_k = 1/\sigma_k^2$ and the weights collected in $\mathbf{W} = \operatorname{diag}(w_k)$, the same matrix is $\boldsymbol{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$, with $\mathbf{J}$ there the Jacobian of the unnormalized rows. For rows linear in the state the normal equations reach $\hat{\mathbf{z}}$ in one step from any $\mathbf{z}_0$, and equations (3.25) and (3.26) are the exact posterior mean and covariance. For nonlinear rows the same step is the Gauss–Newton step, repeated from the new point until it vanishes, and the result is a local Gaussian approximation, the Laplace approximation, whose derivation and curvature term Chapter 4 gives, with the three-entry hall as its example; Section 3.5.6 reads the two-entry hall of Chapter 1 in the terms of this section.

**Precisions add.** Split the rows into prior rows $\mathbf{L}_0(\mathbf{z} - \boldsymbol{\mu}_0)$, whose Jacobian $\mathbf{L}_0$ satisfies $\mathbf{L}_0^\mathsf{T}\mathbf{L}_0 = \boldsymbol{\Sigma}_0^{-1}$, and reading rows $\mathbf{R}^{-1/2}(\mathbf{A}\mathbf{z} - \mathbf{y})$, where the matrix $\mathbf{A}$ maps the state to the quantities read and the diagonal matrix $\mathbf{R}$ holds the variances of the readings. The Jacobian of the reading rows is $\mathbf{J}_y = \mathbf{R}^{-1/2}\mathbf{A}$; stacking the two blocks of rows stacks their Jacobians, and $\mathbf{J}^\mathsf{T}\mathbf{J}$ is the sum of the two blocks' products,

$$\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Sigma}_0^{-1} + \mathbf{J}_y^\mathsf{T}\mathbf{J}_y = \boldsymbol{\Sigma}_0^{-1} + \mathbf{A}^\mathsf{T}\mathbf{R}^{-1}\mathbf{A}. \tag{3.27}$$

Starting the normal equations at $\mathbf{z}_0 = \boldsymbol{\mu}_0$, where the prior rows vanish, gives $\mathbf{J}^\mathsf{T} r(\boldsymbol{\mu}_0) = \mathbf{A}^\mathsf{T}\mathbf{R}^{-1}(\mathbf{A}\boldsymbol{\mu}_0 - \mathbf{y})$, and equation (3.25) gives

$$\hat{\mathbf{z}} = \boldsymbol{\mu}_0 + \boldsymbol{\Sigma}\,\mathbf{A}^\mathsf{T}\mathbf{R}^{-1}(\mathbf{y} - \mathbf{A}\boldsymbol{\mu}_0). \tag{3.28}$$

The posterior information is the prior information plus the information of the readings, and the posterior mean is moved from the prior mean toward the readings in proportion to their information.

**The conditional of a joint Gaussian.** The same posterior follows from the joint distribution of the state and the readings. Before the readings, $\mathbf{z} \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ and $\mathbf{y} = \mathbf{A}\mathbf{z} + \boldsymbol{\epsilon}$ with $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ independent of the state, so the stacked vector of $\mathbf{z}$ and $\mathbf{y}$ is Gaussian with means $\boldsymbol{\mu}_0$ and $\mathbf{A}\boldsymbol{\mu}_0$ and covariance blocks

$$\boldsymbol{\Sigma}_{zz} = \boldsymbol{\Sigma}_0, \qquad \boldsymbol{\Sigma}_{zy} = \boldsymbol{\Sigma}_0\mathbf{A}^\mathsf{T} = \boldsymbol{\Sigma}_{yz}^\mathsf{T}, \qquad \boldsymbol{\Sigma}_{yy} = \mathbf{A}\boldsymbol{\Sigma}_0\mathbf{A}^\mathsf{T} + \mathbf{R}.$$

Let $\mathbf{P}$ be the inverse of the joint covariance and $\mathbf{S} = \boldsymbol{\Sigma}_{zz} - \boldsymbol{\Sigma}_{zy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yz}$ the *Schur complement* of $\boldsymbol{\Sigma}_{yy}$. The block-inversion formula, checked by multiplying back, gives $\mathbf{P}_{zz} = \mathbf{S}^{-1}$ and $\mathbf{P}_{zy} = -\mathbf{S}^{-1}\boldsymbol{\Sigma}_{zy}\boldsymbol{\Sigma}_{yy}^{-1}$. With the readings fixed, the exponent of the joint density, times $-1$, is a quadratic in the state,

$$\tfrac{1}{2}(\mathbf{z}-\boldsymbol{\mu}_0)^\mathsf{T}\mathbf{P}_{zz}(\mathbf{z}-\boldsymbol{\mu}_0) + (\mathbf{z}-\boldsymbol{\mu}_0)^\mathsf{T}\mathbf{P}_{zy}(\mathbf{y}-\mathbf{A}\boldsymbol{\mu}_0) + \text{constant} = \tfrac{1}{2}(\mathbf{z}-\mathbf{m})^\mathsf{T}\mathbf{S}^{-1}(\mathbf{z}-\mathbf{m}) + \text{constant},$$

with $\mathbf{m} = \boldsymbol{\mu}_0 + \boldsymbol{\Sigma}_{zy}\boldsymbol{\Sigma}_{yy}^{-1}(\mathbf{y} - \mathbf{A}\boldsymbol{\mu}_0)$. The conditional density is the joint density divided by $p(\mathbf{y})$, which is free of the state, so

$$\mathbf{z} \mid \mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}_0 + \boldsymbol{\Upsilon}(\mathbf{y} - \mathbf{A}\boldsymbol{\mu}_0),\ \boldsymbol{\Sigma}_0 - \boldsymbol{\Upsilon}\mathbf{A}\boldsymbol{\Sigma}_0), \qquad \boldsymbol{\Upsilon} = \boldsymbol{\Sigma}_0\mathbf{A}^\mathsf{T}(\mathbf{A}\boldsymbol{\Sigma}_0\mathbf{A}^\mathsf{T} + \mathbf{R})^{-1}, \tag{3.29}$$

since $\boldsymbol{\Sigma}_{zy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yz} = \boldsymbol{\Upsilon}\mathbf{A}\boldsymbol{\Sigma}_0$. This is the correction form: the prior mean is moved toward the readings by the gain $\boldsymbol{\Upsilon}$ times the innovation $\mathbf{y} - \mathbf{A}\boldsymbol{\mu}_0$, and the covariance is the Schur complement. The Kalman literature writes the gain $\mathbf{K}$; here it is $\boldsymbol{\Upsilon}$, clear of a controller's gains, and Chapter 9

meets this form again as the filter's update. The correction form gives the same posterior as equations (3.27) and (3.28), by the matrix inversion identity $(\mathbf{\Sigma}_0^{-1}+\mathbf{A}^\mathsf{T}\mathbf{R}^{-1}\mathbf{A})^{-1} = \mathbf{\Sigma}_0-\mathbf{\Upsilon}\mathbf{A}\mathbf{\Sigma}_0$; the information form inverts a matrix of the size of the state, the correction form one of the size of the readings.

*Remark* 3.1. For one entry and one reading, $\mathbf{A} = 1$, $\mathbf{\Sigma}_0 = \sigma_0^2$ and $\mathbf{R} = \sigma_T^2$, and the three equations reduce to the scalar rules $1/\sigma_{\text{post}}^2 = 1/\sigma_0^2 + 1/\sigma_T^2$, $\mu_{\text{post}} = \sigma_{\text{post}}^2(\mu_0/\sigma_0^2 + y_T/\sigma_T^2)$ and $\mu_{\text{post}} = \mu_0 + \upsilon(y_T - \mu_0)$ with $\upsilon = \sigma_0^2/(\sigma_0^2 + \sigma_T^2)$: precisions add, and the posterior mean is the precision-weighted average of the prior mean and the reading.

### 3.5.4 The numbers and what changed

For the hall the state is the one number $T$, and the two rows of equation (3.23) have the slopes

$$j_1 = \frac{1}{\sigma_0} = 0.5\,\mathrm{K}^{-1}, \qquad j_2 = \frac{1}{\sigma_T} = 2\,\mathrm{K}^{-1}.$$

The row terms are $\mathcal{L}_k = \frac{1}{2}r_k^2$, with the slopes $\mathcal{L}_k' = j_k r_k$ and the curvatures $\mathcal{L}_k'' = j_k^2$ in $T$, the primes denoting $\mathrm{d}/\mathrm{d}T$; slopes and curvatures add over the rows,

$$\mathcal{L}' = j_1 r_1 + j_2 r_2, \qquad \mathcal{L}'' = j_1^2 + j_2^2.$$

Because the rows are linear, the curvature has no further term and is the same at every $T$; it is the information

$$\Lambda = j_1^2 + j_2^2 = 0.5^2 + 2^2 = 0.25 + 4 = 4.25\,\mathrm{K}^{-2},$$

the prior precision $1/\sigma_0^2$ plus the reading precision $1/\sigma_T^2$, as equation (3.27) states in the scalar form of the remark above, with the prior variance $\sigma_0^2 = 4\,\mathrm{K}^2$ and the reading variance $\sigma_T^2 = 0.25\,\mathrm{K}^2$. It is the one-entry case of $\mathbf{J}^\mathsf{T}\mathbf{J}$, and equation (3.26) gives the posterior variance as its reciprocal,

$$\sigma_{\text{post}}^2 = \frac{1}{\Lambda} = \frac{1}{4.25} = 0.2353\,\mathrm{K}^2, \qquad \sigma_{\text{post}} = 0.4851\,\mathrm{K}.$$

With one entry the normal equations (3.25) are the one equation $\Lambda\,\Delta T = -\mathcal{L}'$, the Newton step $\Delta T = -\mathcal{L}'/\mathcal{L}''$. They start at the prior mean, $T_0 = 26°\mathrm{C}$, where the prior row is 0 and the reading row $(26-27)/0.5 = -2$, so that $\mathcal{L}' = 0.5\times 0 + 2\times(-2) = -4\,\mathrm{K}^{-1}$, and the step is $\Delta T = 4/4.25 = 0.9412\,\mathrm{K}$. The iterations are

| $k$ [–] | $T$ [°C] | $\mathcal{L}(T)$ [–] | $\mathcal{L}_1'$ [K$^{-1}$] | $\mathcal{L}_2'$ [K$^{-1}$] | $\mathcal{L}'$ [K$^{-1}$] | $\mathcal{L}''$ [K$^{-2}$] | $\Delta T$ [K] |
|---|---|---|---|---|---|---|---|
| 0 | 26 | 2.0000 | 0 | −4 | −4 | 4.25 | 0.9412 |
| 1 | 26.9412 | 0.1176 | 0.2353 | −0.2353 | 0 | 4.25 | 0 |

At the second point the residuals are $r_1 = 0.4706$ and $r_2 = -0.1176$, the slope is $\mathcal{L}' = 0.5\times 0.4706 + 2\times(-0.1176) = 0$, and the step vanishes: the rows are linear, and one step reaches the minimum. The correction form (3.29) confirms both moments: the variance of the reading before it arrives is $\sigma_0^2 + \sigma_T^2 = 4 + 0.25 = 4.25\,\mathrm{K}^2$, $\upsilon = 4/4.25 = 0.9412$, the mean $26 + 0.9412\times(27-26) = 26.9412°\mathrm{C}$ and the Schur complement $4 - 0.9412\times 4 = 0.2353\,\mathrm{K}^2$. The answer is the minimum $\hat{T} = 26.9412°\mathrm{C}$, written $\mu_{\text{post}}$ below, with the standard deviation $\sigma_{\text{post}}$:

$$T \mid y_T \sim \mathcal{N}(26.9412, 0.4851^2). \tag{3.30}$$

Figure 3.6(a) draws the update. The horizontal axis is the temperature and the vertical axis the probability density. The gray curve is the prior, the dashed blue curve the likelihood of the reading as a function of $T$ (normalized to integrate to one), and the green curve the posterior. The posterior mean moved from 26 to 26.94°C, and the standard deviation fell from 2 to 0.49 K.

The reading was more precise than the prior, so it received the greater weight, $\upsilon = 0.94$; but the posterior mean is not exactly 27°C, because the prior also carries information.

Table 3.7 shows the trend for sensors of four precisions, all reading 27°C against the same prior. The first column is the sensor's standard deviation, the second the weight $\upsilon$ of the reading, and the remaining columns the posterior mean, the posterior standard deviation and the exceedance probability derived in Section 3.5.5. Each row is the same computation with the slopes $j_1 = 1/\sigma_0$ and $j_2 = 1/\sigma_T$, so that $\sigma^2_{\text{post}} = 1/\Lambda = 1/(0.25 + 1/\sigma_T^2)$. The second row is the case worked above. A sensor as uncertain as the prior, $\sigma_T = 2\,\text{K}$, receives half the weight and leaves the posterior standard deviation at $2/\sqrt{2} = 1.41\,\text{K}$.

Table 3.7: The hall-air update for four sensor precisions: prior $\mathcal{N}(26, 2^2)$, reading $y_T = 27$°C, equations (3.27) and (3.29), and the exceedance probability of $T_{\max} = 28$°C derived in Section 3.5.5.

| $\sigma_T$ [K] | weight $\upsilon$ [–] | $\mu_{\text{post}}$ [°C] | $\sigma_{\text{post}}$ [K] | $\Pr(T > T_{\max} \mid y_T)$ [–] |
|---|---|---|---|---|
| 0.25 | 0.9846 | 26.9846 | 0.2481 | $2 \times 10^{-5}$ |
| 0.5 | 0.9412 | 26.9412 | 0.4851 | 0.0145 |
| 1.0 | 0.8000 | 26.8000 | 0.8944 | 0.0899 |
| 2.0 | 0.5000 | 26.5000 | 1.4142 | 0.1444 |

### 3.5.5 The threshold question

Suppose the hall-air limit of this illustration is $T_{\max} = 28$°C, and the question is the probability that the hall exceeds it, $\Pr(T > T_{\max} \mid y_T)$. Since the posterior is Gaussian, standardizing with the mean $\mu_{\text{post}} = \hat{T}$ and the standard deviation $\sigma_{\text{post}} = 1/\sqrt{\Lambda}$ and using the standard normal cumulative distribution function $\Phi$,

$$\begin{aligned} \Pr(T > T_{\max} \mid y_T) &= 1 - \Phi\left(\frac{T_{\max} - \mu_{\text{post}}}{\sigma_{\text{post}}}\right) \\ &= 1 - \Phi\left(\frac{28 - 26.9412}{0.4851}\right) = 1 - \Phi(2.183) = 0.0145, \end{aligned} \tag{3.31}$$

or about 1.45 %. Before the reading the same calculation with the prior gives $\Pr(T > 28\text{°C}) = 1 - \Phi\big((28 - 26)/2\big) = 1 - \Phi(1) = 0.1587$. The estimated exceedance probability fell from 15.9 % to 1.45 %, for the same fixed operating condition and the assumed Gaussian measurement model. The rows are linear, so this normal tail is the exact posterior's: a direct quadrature of the product (3.22) also gives 0.0145, and there is no gap to set beside it, as there is on the curved physics row of the next subsection. Figure 3.6(b) shades the two tails: the axes are those of panel (a), the dashed line is the limit, the gray area to its right under the prior is 0.159, and the green area under the posterior is 0.0145.

The physical hall did not necessarily change when the reading arrived; the information about it changed, and that is the mechanism of Bayesian updating. The calculation also walks the method of *How to read this book* on its smallest example. The plant graph of Figure 3.5(a) is step 0. The question $\Pr(T > T_{\max} \mid y_T)$, unknown on the left and reading on the right, is step 1. Attaching the prior and the reading as the two rows of equation (3.23) is step 2. The posterior, derived through the normal equations and the variance as the reciprocal of the information, equations (3.25) and (3.26) with one entry, and computed by the one Newton step $-\mathcal{L}'/\mathcal{L}''$ with $\Lambda = 4.25\,\text{K}^{-2}$, is step 3. Reading the normal tail from it is step 4, and checking that tail by the correction form and its Schur complement is step 5, and the sentence "the hall exceeds 28°C with probability 1.45 %, posterior $26.94 \pm 0.49\,\text{K}$" is step 6. Step 7, the next reading, reads

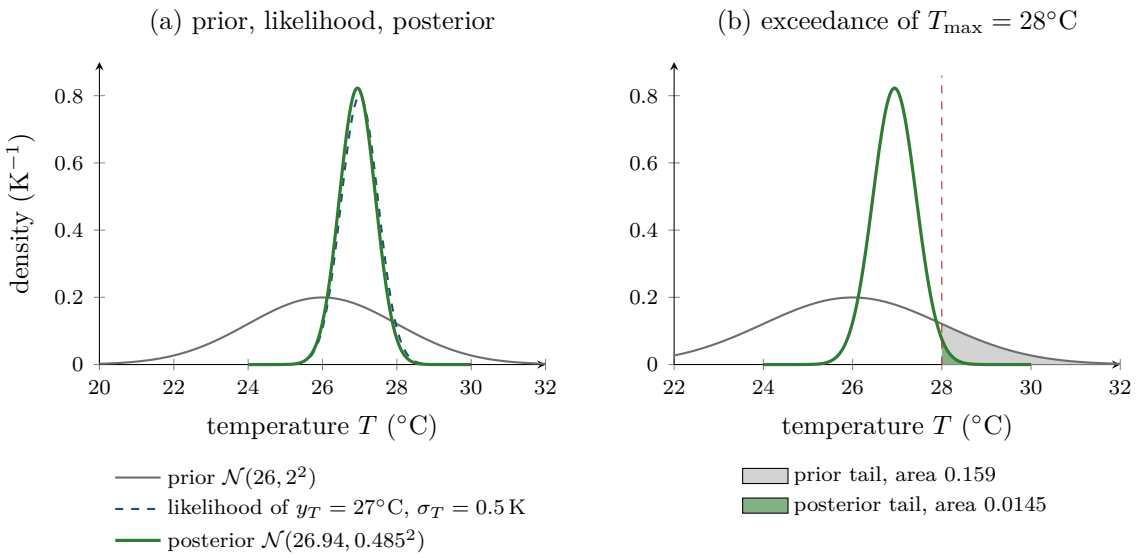



Figure 3.6: The hall-air update of equations (3.27)–(3.30). (a) Prior $\mathcal{N}(26, 2^2)$, likelihood of the reading $y_T = 27°\mathrm{C}$ with $\sigma_T = 0.5\,\mathrm{K}$ drawn as the density $\mathcal{N}(27, 0.5^2)$ in $T$, and posterior $\mathcal{N}(26.9412, 0.4851^2)$. (b) The probability that the hall exceeds $T_{\max} = 28°\mathrm{C}$ (dashed): $1-\Phi(1) = 0.159$ under the prior and $1-\Phi(2.183) = 0.0145$ under the posterior (equation (3.31)).

Table 3.7: if 1.45 % were too high a residual doubt, a sensor with $\sigma_T = 0.25\,\mathrm{K}$ would, for a reading near the same value, reduce it to the order of $10^{-5}$; Chapter 11 asks which sensors can resolve what the controller needs. The posterior is inferred first, and the engineer's question is then answered from it:

$$\underbrace{p(T \mid y_T)}_{\text{inference}} \quad \longrightarrow \quad \underbrace{\Pr(T > T_{\max} \mid y_T)}_{\text{threshold question}}.$$

### 3.5.6 The two-entry hall and its physics row

The reading of the one-entry hall informed only the entry it measured; Section 3.5.1 claimed that a reading also reaches the conductance through the physics row. The hall of Chapter 1 shows it. Its state is $\mathbf{z} = (T,\ g)^{\mathsf{T}}$, and its three rows, equation (1.10), are the physics row with tolerance $\sigma_p = 0.01\,\mathrm{K}$, the reading row with $\sigma_T = 0.5\,\mathrm{K}$ and the handbook prior row on $g$ with standard deviation 0.3; Chapter 1 solves their normal equations by Gauss–Newton from (23°C, 0) (Table 1.4) and reaches $\hat{\mathbf{z}} = (23.070°\mathrm{C},\ -0.0139)^{\mathsf{T}}$. Equation (3.27) holds row by row: with $\mathbf{J}_k$ the $k$-th row of $\mathbf{J}$, $\mathbf{J}^{\mathsf{T}}\mathbf{J} = \sum_k \mathbf{J}_k^{\mathsf{T}}\mathbf{J}_k$, each row contributing the outer product of its own sensitivities. At $\hat{\mathbf{z}}$ the rows of $\mathbf{J}$ are (100, 507.0), (2, 0) and (0, 3.333), and

$$\mathbf{J}^{\mathsf{T}}\mathbf{J} = \underbrace{\begin{bmatrix} 10\,000 & 50\,701 \\ 50\,701 & 257\,056 \end{bmatrix}}_{\text{physics}} + \underbrace{\begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}}_{\text{reading}} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 11.11 \end{bmatrix}}_{\text{prior}} = \begin{bmatrix} 10\,004 & 50\,701 \\ 50\,701 & 257\,067 \end{bmatrix}.$$

The reading's and the prior's terms are diagonal, because each of those rows carries one entry. The only off-diagonal entry, 50 701, comes from the physics row, the one row with sensitivities in both columns: this is equation (3.20) in vector form, the reading informing $T$ and the physics row carrying that information to $g$. Without the reading row, the physics and prior terms alone invert to $\sqrt{\Sigma_{22}} = 0.3$, the prior spread; with it, the covariance of Chapter 1, equation (1.21), has $\sqrt{\Sigma_{11}} = 0.481\,\mathrm{K}$, $\sqrt{\Sigma_{22}} = 0.0948$ and correlation $-0.9998$. The reading has cut the conductance spread from 0.3 to 0.0948 although its own row has a zero in the column of $g$, and the posterior lies along the physics curve, on the short arc of the state plane of Figure 1.5 inside the reading

band. Because the physics row is nonlinear, these spreads include the curvature term that Chapter 4 derives; $(\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ alone gives 0.475 K and 0.094. The exact posterior of these rows has the standard deviation 0.4813 K against the Laplace 0.4806 K, a gap of 0.0007 K, and the mean 23.0300°C against the mode 23.0701°C.

### 3.5.7 Dependence inside and outside the physics graph

The common cause of Section 3.4.4 cost a factor of nine because it was left out of the product. A plant graph built on conservation relations does not leave out a shared cause that has a physical path: the cause is an entry of the state, and every row that depends on it carries it. The two-entry hall above shows this. The reading and the conductance share no row, yet the physics row couples them, and the posterior correlation of $T$ and $g$, $-0.9998$, comes out of $\mathbf{J}^\mathsf{T}\mathbf{J}$ without being assumed. Had the pump been a node of the plant graph, with a row for each loop's trip given its state, the joint trip probability (3.17) would have followed from the rows, and the independent product of Section 3.4.5 could not have been formed. Conditional independence given the state is then the statement of Section 3.4.6 with the common cause known: it holds because the cause is in the state.

What the graph does not supply is the independence of the errors. Equation (3.20) assumes that a reading's error is independent of the inputs and, when there are several readings, of the other readings' errors, so that the posterior (3.24) is a product over rows. No conservation relation states that, and none checks it directly. Suppose the hall of Section 3.5.4 carries two sensors from one calibration batch, each with $\sigma_T = 0.5\,\mathrm{K}$, both reading 27°C, with errors correlated through the offset they share, $\mathrm{corr}(\epsilon_1, \epsilon_2) = 0.5$. Treated as independent rows, the pair adds the precision $2/\sigma_T^2 = 8\,\mathrm{K}^{-2}$ to the prior's 0.25. With the correlation, equation (3.27) holds with $\mathbf{R}$ a full matrix, its off-diagonal entry $0.5\,\sigma_T^2$ and its rows whitened by $\mathbf{R}^{-1/2}$, and the pair adds only

$$\mathbf{A}^\mathsf{T}\mathbf{R}^{-1}\mathbf{A} = \frac{2}{\sigma_T^2\,(1 + 0.5)} = 5.33\,\mathrm{K}^{-2}, \qquad \mathbf{A} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

the precision of the pair's average, whose error has the variance $\sigma_T^2\,(1 + 0.5)/2$. Table 3.8 compares the two treatments with the single sensor of Table 3.7. Its columns are the precision the readings add, the posterior mean, the posterior standard deviation and the probability of exceeding $T_{\max} = 28$°C.

Table 3.8: Two hall sensors of $\sigma_T = 0.5\,\mathrm{K}$, both reading 27°C, against the prior $\mathcal{N}(26, 2^2)$ and the limit $T_{\max} = 28$°C: the pair treated as independent, the pair with $\mathrm{corr}(\epsilon_1, \epsilon_2) = 0.5$, and one sensor, which is also the limit of a correlation of one.

| Readings | precision added [$\mathrm{K}^{-2}$] | mean [°C] | $\sigma_{\mathrm{post}}$ [K] | $\Pr(T > T_{\max} \mid \mathbf{y})$ [–] |
|---|---|---|---|---|
| two, treated as independent | 8.00 | 26.9697 | 0.3482 | 0.00154 |
| two, correlation 0.5 | 5.33 | 26.9552 | 0.4232 | 0.00678 |
| one sensor | 4.00 | 26.9412 | 0.4851 | 0.0145 |

The independent treatment reports an exceedance probability of 0.00154 against the 0.00678 the correlated pair supports, 4.4 times too small, and a posterior standard deviation of 0.348 K against 0.423 K. The error has the direction of Section 3.4.5's: a posterior narrower than the evidence supports, and a limit that looks safer than it is. At a correlation of one the second sensor adds nothing, and the posterior is that of one sensor.

The shared offset is a common cause, and the remedy of Section 3.4.6 applies: put it in the state. A bias entry $b \sim \mathcal{N}(0,\ 0.5\,\sigma_T^2)$ carried by both reading rows, with independent remainders of variance $0.5\,\sigma_T^2$, makes the two rows conditionally independent given $(T, b)$, and the state $(T, b)$ gives the posterior of $T$ of the correlated row of Table 3.8 to every printed digit. The posterior of $b$, however, has the standard deviation 0.348 K against its prior's 0.354 K, with $\mathrm{corr}(T, b \mid \mathbf{y}) = -0.81$. The readings barely inform it: a shared offset moves both readings together, as a change of $T$ does, and only the prior on $b$ bounds it.

Inside a loop the entry has a second consequence. A controller holds its reading at the set point, not the quantity the reading is meant to measure. An offset on that sensor is therefore not scatter that a longer record averages away, since every window carries it, as both rows above carry $b$. It does not show in the loop's error either, which compares the set point with the reading and is blind to the gap between the reading and the air: the trend stays on its set point with the offset present. The offset is one more entry of the state, and this chapter names it and leaves it there. Chapter 6 and Chapter 26 derive what it does to the protected quantity.

The same assumption hides in four other places a physics graph does not reach: a prior shared by several entries, such as two conductances taken from one handbook; a physics tolerance $\sigma_p$ treated as independent from row to row where the model's error is systematic; repeated readings averaged as if their errors were independent when they are correlated in time; and a path of the conserved quantity that the graph omits. Each makes the posterior narrower than the evidence supports. What the graph adds is redundancy. When more rows bear on the state than it has entries, a misfit of the rows taken together, a reading far from what the other rows predict, or a comparison of drawings with and without the suspected cause can reveal the dependence: Chapter 26 takes up the misfit and the reading far from its prediction, Chapter 27 the comparison of drawings on a running loop, and the two-hall plant of Chapter 18 carries the shared chilled water as an entry of the state. No test reveals a dependence that moves the readings along a direction the physics also allows, as the shared offset moves them along $T$.

**Principle 3.3.** *A physics graph carries dependence through shared physical quantities as entries of the state, so the independent product cannot be formed for them. The independence it still assumes is that of the errors: it must be stated, tested where the rows are redundant, and, where it fails, removed by putting the shared cause in the state.*

## 3.6 Predictive variance and the covariance of a joint posterior

Two variances complete the chapter: the law of total variance, which separates the randomness of future trips from the uncertainty of their rate, and the covariance of a posterior over a joint state, which decides how uncertain a derived quantity is and why the shape of the posterior must be read rather than its marginals.

### 3.6.1 The law of total variance

Suppose the number $K_\tau$ of the loop's trips during a future period of length $\tau$ is Poisson given the rate, $K_\tau \mid \lambda \sim \mathrm{Poisson}(\lambda\tau)$, and the rate itself is uncertain, with posterior $p(\lambda \mid \mathbf{y})$. The predictive variance of $K_\tau$ is wanted. Two facts give it. The law of total expectation, which is the law of total probability applied inside an expectation, states that $\mathbb{E}[K_\tau \mid \mathbf{y}] = \mathbb{E}\big[\mathbb{E}[K_\tau \mid \lambda] \mid \mathbf{y}\big]$, the conditional mean averaged over the rate; and the variance of any quantity is the mean of its square minus the square of its mean. Then

$$\mathrm{Var}(K_\tau \mid \mathbf{y}) = \mathbb{E}\big[\mathbb{E}[K_\tau^2 \mid \lambda] \mid \mathbf{y}\big] - \big(\mathbb{E}\big[\mathbb{E}[K_\tau \mid \lambda] \mid \mathbf{y}\big]\big)^2,$$

and writing the conditional second moment as $\mathbb{E}[K_\tau^2 \mid \lambda] = \mathrm{Var}(K_\tau \mid \lambda) + \mathbb{E}[K_\tau \mid \lambda]^2$,

$$\mathrm{Var}(K_\tau \mid \mathbf{y}) = \mathbb{E}\big[\mathrm{Var}(K_\tau \mid \lambda) \mid \mathbf{y}\big] + \Big\{\mathbb{E}\big[\mathbb{E}[K_\tau \mid \lambda]^2 \mid \mathbf{y}\big] - \big(\mathbb{E}\big[\mathbb{E}[K_\tau \mid \lambda] \mid \mathbf{y}\big]\big)^2\Big\}.$$

The brace is the variance of the conditional mean, which gives the *law of total variance*,

$$\operatorname{Var}(K_\tau \mid \mathbf{y}) = \mathbb{E}\big[\operatorname{Var}(K_\tau \mid \lambda) \mid \mathbf{y}\big] + \operatorname{Var}\big(\mathbb{E}[K_\tau \mid \lambda] \mid \mathbf{y}\big). \tag{3.32}$$

The first term is the expected variability of the trip count when the rate is specified, the aleatory part. The second is the additional predictive variability because the rate is uncertain, the epistemic part. For a Poisson count the mean and the variance are equal, $\mathbb{E}[K_\tau \mid \lambda] = \operatorname{Var}(K_\tau \mid \lambda) = \lambda\tau$, so that $\operatorname{Var}(K_\tau \mid \mathbf{y}) = \mathbb{E}[\lambda\tau \mid \mathbf{y}] + \operatorname{Var}(\lambda\tau \mid \mathbf{y})$, and since $\tau$ is fixed,

$$\operatorname{Var}(K_\tau \mid \mathbf{y}) = \tau\,\mathbb{E}[\lambda \mid \mathbf{y}] + \tau^2 \operatorname{Var}(\lambda \mid \mathbf{y}). \tag{3.33}$$

The predictive variance has been separated into a process term and a rate term.

### 3.6.2 Trip counts over the next run

Take the posterior of the loop's trip rate from the state, the normal of the log-rate of equation (2.41), with $\hat{\omega} = \ln(8/3)$ and $\sigma_\omega = 0.5$, so that $\lambda$ is lognormal about $\hat{\lambda} = 4/1500\,\mathrm{h}$. The standard lognormal moments give

$$\mathbb{E}[\lambda \mid \mathbf{y}] = \hat{\lambda}\,\mathrm{e}^{\sigma_\omega^2/2} = 3.022 \times 10^{-3}\,\mathrm{h}^{-1}, \qquad \operatorname{Var}(\lambda \mid \mathbf{y}) = \hat{\lambda}^2\,\mathrm{e}^{\sigma_\omega^2}\big(\mathrm{e}^{\sigma_\omega^2} - 1\big) = 2.593 \times 10^{-6}\,\mathrm{h}^{-2}.$$

Over the next night run, $\tau = 12\,\mathrm{h}$, the expected trip count is $\mathbb{E}[K_\tau \mid \mathbf{y}] = 12 \times 3.022 \times 10^{-3} = 0.0363$, and by equation (3.33) the predictive variance is

$$\operatorname{Var}(K_\tau \mid \mathbf{y}) = 0.0363 + 12^2 \times 2.593 \times 10^{-6} = 0.0363 + 0.0004 = 0.0366.$$

The process term is 99.0 % of the predictive variance and the rate term 1.02 %. The count 0.0363 is a mean, not the probability of a trip: a night with two trips counts twice in the mean, so the probability of at least one, 0.0354 from the same normal, lies below it.

The exact posterior $\text{Gamma}(a, b) = \text{Gamma}(4, 1500\,\mathrm{h})$ stands beside these, with the standard gamma moments $\mathbb{E}[\lambda \mid \mathbf{y}] = a/b = 2.667 \times 10^{-3}\,\mathrm{h}^{-1}$, one trip in about 375 h, and $\operatorname{Var}(\lambda \mid \mathbf{y}) = a/b^2 = 1.778 \times 10^{-6}\,\mathrm{h}^{-2}$. The same two terms are

$$\operatorname{Var}(K_\tau \mid \mathbf{y}) = 12\left(\frac{4}{1500}\right) + 12^2\left(\frac{4}{1500^2}\right) = 0.032000 + 0.000256 = 0.032256,$$

the mean 0.0320 against the reported 0.0363, a gap of 0.0043, and the variance 0.0323 against 0.0366; the probability of a trip, 0.0314, again lies below its mean. The share of the rate term under the gamma is

$$\frac{\tau^2 a/b^2}{\tau a/b + \tau^2 a/b^2} = \frac{\tau}{b + \tau},$$

which for $\tau = 12\,\mathrm{h}$ and $b = 1500\,\mathrm{h}$ is $12/1512 = 0.0079$, against the reported 0.0102. The gamma's result can be checked against its full predictive distribution. A Poisson count whose mean is averaged over a gamma rate has, by a standard result, the negative binomial distribution with shape $a$ and success probability $\pi = b/(b + \tau)$, whose mean is $a(1 - \pi)/\pi = a\tau/b$ and whose variance is $a(1 - \pi)/\pi^2 = a\tau/b + a\tau^2/b^2$: exactly the two terms of equation (3.33). For $\pi = 1500/1512 = 0.99206$ they are 0.032000 and 0.032256.

Table 3.9 extends the calculation to longer periods. The first column is the period; the next four are the reported values from the state, the expected count, which is also the process term, the rate term, the total and the share of the rate term; and the last three the gamma's mean, variance and share. Over a run almost all of the predictive variance comes from the random trip process, and a longer trip log, which shrinks only the rate term, would barely change the forecast of the next run's trips. The two terms are equal at about 1165 h (at 1500 h under the gamma, the logged 1000 h plus the prior's 500 h), and over 3000 h the rate term is 0.720 of the total (0.667 under the gamma). The gaps grow with the period, because the lognormal's mean rate lies above the gamma's and its upper tail is heavier. Whether a longer log will meaningfully improve a trip forecast depends on this split: more evidence shrinks the second term, and the first remains as long as the future count is modeled as a Poisson process.

Table 3.9: Predictive mean and variance of the loop's trip count over a future period $\tau$, equation (3.33): reported from the normal of the log-rate, with the posterior Gamma$(4, 1500\,\mathrm{h})$ beside it; the first row is one night run.

| $\tau$ [h] | mean = process term [–] | rate term [–] | variance [–] | rate share [–] | gamma mean [–] | gamma variance [–] | gamma share [–] |
|---|---|---|---|---|---|---|---|
| 12 | 0.0363 | 0.0004 | 0.0366 | 0.0102 | 0.0320 | 0.0323 | 0.0079 |
| 100 | 0.3022 | 0.0259 | 0.3281 | 0.0790 | 0.2667 | 0.2844 | 0.0625 |
| 500 | 1.5109 | 0.6483 | 2.1592 | 0.3003 | 1.3333 | 1.7778 | 0.2500 |
| 1500 | 4.5326 | 5.8351 | 10.3677 | 0.5628 | 4.0000 | 8.0000 | 0.5000 |
| 3000 | 9.0652 | 23.3405 | 32.4057 | 0.7203 | 8.0000 | 24.0000 | 0.6667 |

### 3.6.3 One posterior over a joint state

The examples so far had one unknown at a time. A plant graph carries many at once. Suppose the state vector holds the air temperatures of two halls, the log-conductance of an air path and the trip rates of the two loops,

$$\mathbf{z} = \begin{bmatrix} T_A & T_B & g & \lambda_A & \lambda_B \end{bmatrix}^{\mathsf{T}},$$

and the evidence $\mathbf{y}$ gathers temperature readings, the trip logs and stroke tests. Everything known after the evidence is the joint posterior $p(\mathbf{z} \mid \mathbf{y}, M)$, where $M$ names the model: the graph, its rows and its priors. Different questions are asked of this same distribution:

*State estimation,* $p(T_A \mid \mathbf{y}, M)$: the marginal of one entry.

*Threshold probability,* $\Pr(T_B > T_{\max} \mid \mathbf{y}, M)$: a tail of another marginal.

*Parameter comparison,* $\Pr(\lambda_A > \lambda_B \mid \mathbf{y}, M)$: a tail of the derived quantity $\lambda_A - \lambda_B$, which depends on the joint distribution of the two rates.

*Expected derived quantity,* $\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}, M]$, such as the expected number of trips in the two loops over the next run: the average of a function of the state.

*Prediction,* $\Pr(\text{no trip in } \tau \mid \mathbf{y}, M)$: a posterior predictive probability, equation (2.41) with the rate taken from the state.

The methods of computation differ, but every one of them depends on a consistent probability model of the physical plant and the evidence, and none needs a separate inference.

### 3.6.4 The posterior covariance and the variance budget

When the posterior can be approximated locally by a multivariate Gaussian, $\mathbf{z} \mid \mathbf{y}, M \approx \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, the covariance matrix $\boldsymbol{\Sigma}$ describes the uncertainty of each entry and how the entries vary together. The approximation takes the most probable state as the mean and the inverse curvature of the objective there as the covariance. For rows linear in the state it is exact, $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ by Section 3.5.3; for nonlinear rows it is the Laplace approximation, which Chapter 4 derives with its curvature term. For two entries,

$$\boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}, \qquad \operatorname{corr}(z_1, z_2 \mid \mathbf{y}) = \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\Sigma_{22}}}.$$

The diagonal entries are the marginal variances, $\Sigma_{11} = \operatorname{Var}(z_1 \mid \mathbf{y})$; the off-diagonal entry is the covariance, and $\operatorname{corr}(z_1, z_2 \mid \mathbf{y})$ the correlation. The off-diagonal entries matter because physical

relations and shared observations introduce posterior dependence between unknowns; they are not zero merely because the priors of the entries were independent. The hall of Section 3.5.6 is one case. Another carries the supply air as an entry: the state $\mathbf{z} = (T,\ T_c,\ g)^{\mathsf{T}}$ for the illustrative hall with a weak air path of Figure 3.2(c), with $Q = 30\,\mathrm{kW}$, $G_{\mathrm{ref}} = 3\,\mathrm{kW/K}$, independent priors $T_c \sim \mathcal{N}(20, 1^2)$ and $g \sim \mathcal{N}(0, 0.05^2)$ (the $0.15\,\mathrm{kW/K}$ of the figure), a reading with $\sigma_T = 0.5\,\mathrm{K}$ equal to the prior prediction 30°C, and the physics row $(T - T_c - 10\,\mathrm{K}\ \mathrm{e}^{-g})/\sigma_p$ with $\sigma_p = 0.01\,\mathrm{K}$. At the prior means every residual is zero and the rows of $\mathbf{J}$ are $(1,\ -1,\ 10)/\sigma_p$ for the physics, $(1,\ 0,\ 0)/\sigma_T$ for the reading, and $(0,\ 1,\ 0)$ and $(0,\ 0,\ 1)/0.05$ for the priors; equation (3.26) gives

$$\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1} = \begin{bmatrix} 0.2083 & 0.1667 & -0.0042 \\ 0.1667 & 0.3334 & 0.0167 \\ -0.0042 & 0.0167 & 0.0021 \end{bmatrix},$$

and the correlation of $T_c$ and $g$ is $0.0167/\sqrt{0.3334 \times 0.0021} = 0.632$ (unrounded entries; 0.6325 at $\sigma_p = 0.001\,\mathrm{K}$). Without the reading row the same computation gives zero. A single reading leaves $T_c$ and $G$ positively correlated, since a warmer supply air and a larger conductance explain the same reading, and the correlation enters through the physics row, the only row with entries in the columns of both.

Suppose a derived quantity $f(\mathbf{z})$ has the gradient $\mathbf{g} = \partial f/\partial \mathbf{z}$ at $\hat{\mathbf{z}}$, a column with one entry per entry of the state. With $\boldsymbol{\delta} = \mathbf{z} - \hat{\mathbf{z}}$ the deviation of the state from its mean, the first-order change is $f - f(\hat{\mathbf{z}}) \approx \mathbf{g}^{\mathsf{T}}\boldsymbol{\delta}$, a scalar equal to its own transpose, and the covariance is $\boldsymbol{\Sigma} = \mathbb{E}[\boldsymbol{\delta}\boldsymbol{\delta}^{\mathsf{T}}]$, so that

$$\mathrm{Var}(f) \approx \mathbb{E}\big[\mathbf{g}^{\mathsf{T}}\boldsymbol{\delta}\,\boldsymbol{\delta}^{\mathsf{T}}\mathbf{g}\big] = \mathbf{g}^{\mathsf{T}}\,\mathbb{E}\big[\boldsymbol{\delta}\boldsymbol{\delta}^{\mathsf{T}}\big]\,\mathbf{g} = \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}. \tag{3.34}$$

This is the *variance budget* of the output. Written out entry by entry, $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ has one term per entry of the state, the squared sensitivity of the output to the entry times the variance of the entry, plus a cross term $2\,\nabla_k \nabla_l\,\Sigma_{kl}$, with $\nabla_k = \partial f/\partial z_k$, for every pair of entries that move together. For the sum $f = z_1 + z_2$ the gradient is $\mathbf{g} = (1,\ 1)^{\mathsf{T}}$, and the budget is exact because $f$ is linear,

$$\mathrm{Var}(z_1 + z_2) = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \Sigma_{11} + \Sigma_{22} + 2\Sigma_{12}. \tag{3.35}$$

With $\Sigma_{11} = \Sigma_{22} = 1\,\mathrm{K}^2$ (illustrative), the variance of the sum is $0.4\,\mathrm{K}^2$ for $\Sigma_{12} = -0.8\,\mathrm{K}^2$, $2\,\mathrm{K}^2$ for independent entries and $3.6\,\mathrm{K}^2$ for $\Sigma_{12} = 0.8\,\mathrm{K}^2$. A negative covariance makes the output more precisely known than the individual entries suggest; a positive one makes it less.

Table 3.10 works the budget for the hall air of equation (3.19), with the load known, $Q = 30\,\mathrm{kW}$, and the two unknowns of Figure 3.2(c) as the state $\mathbf{z} = (T_c,\ G)^{\mathsf{T}}$: $T_c$ with standard deviation 1 K and $G$ with mean $3\,\mathrm{kW/K}$ and standard deviation $0.15\,\mathrm{kW/K}$. The gradient of $T$ and the covariance are

$$\mathbf{g} = \begin{bmatrix} \partial T/\partial T_c \\ \partial T/\partial G \end{bmatrix} = \begin{bmatrix} 1 \\ -Q/G^2 \end{bmatrix} = \begin{bmatrix} 1 \\ -3.333 \end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.15\,c \\ 0.15\,c & 0.0225 \end{bmatrix}, \qquad c = \mathrm{corr}(T_c, G),$$

with $\mathbf{g}$ in K per unit of each entry and $\boldsymbol{\Sigma}$ in the products of the entries' units, so that $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} = (1.00 + 0.25 - 1.00\,c)\,\mathrm{K}^2$: the supply-air term $1^2 \times 1$, the conductance term $3.333^2 \times 0.0225$ and the cross term $2 \times 1 \times (-3.333) \times 0.15\,c$. The columns of the table are the correlation, the three terms, the variance and the standard deviation of $T$. The two marginals are the same in every row, and the standard deviation of the hall air ranges from 0.67 to 1.43 K with the correlation alone. The budget is first order, since $T$ is curved in $G$; a quadrature over the same joint normal gives 0.6701, 1.1203 and 1.4357 K against the budget's 0.6708, 1.1180 and 1.4318 K, gaps of at most 0.0039 K.

Two consequences follow from equation (3.34). An entry can be very uncertain and still not matter: if the output is insensitive to it, its whole term is near zero, and narrowing its variance

Table 3.10: Variance budget of the hall air $T = T_c + Q/G$, equation (3.34), with $Q = 30\,\mathrm{kW}$, $G$ of mean $3\,\mathrm{kW/K}$ and standard deviation $0.15\,\mathrm{kW/K}$, $T_c$ of standard deviation $1\,\mathrm{K}$, and correlation $c = \mathrm{corr}(T_c, G)$ (illustrative).

| $c$ [–] | supply-air term [K$^2$] | conductance term [K$^2$] | cross term [K$^2$] | Var($T$) [K$^2$] | sd [K] |
|---|---|---|---|---|---|
| $-0.8$ | 1.00 | 0.25 | $+0.80$ | 2.05 | 1.43 |
| 0 | 1.00 | 0.25 | 0 | 1.25 | 1.12 |
| 0.8 | 1.00 | 0.25 | $-0.80$ | 0.45 | 0.67 |

buys nothing. And an entry can be poorly known while the output is well known: when the sensitivities and the covariance have opposite signs, the cross term subtracts from the diagonal terms. A reading is worth requesting when it shrinks a variance $\Sigma_{kk}$ whose sensitivity is large, and worthless when the sensitivity is near zero.

### 3.6.5 Reading the shape of the posterior

The second consequence deserves a worked case, because it is where marginal spreads mislead. Carry the hall air, the load and the conductance as entries, the last two through the logarithms $\ell_Q = \ln(Q/Q_{\mathrm{ref}})$ and $g = \ln(G/G_{\mathrm{ref}})$, each with an independent prior of mean 0 and standard deviation 0.10, that is about 10 %, and with the supply air pinned. The references are chosen so that $Q_{\mathrm{ref}}/G_{\mathrm{ref}} = 10\,\mathrm{K}$, a temperature rise of 10 K at the prior means, and a sensor with $\sigma_T = 0.2\,\mathrm{K}$ reads that temperature, $y_T = T_c + 10\,\mathrm{K}$. The state, its four rows and their $4 \times 3$ Jacobian are

$$\mathbf{z} = \begin{bmatrix} T \\ \ell_Q \\ g \end{bmatrix}, \qquad r(\mathbf{z}) = \begin{bmatrix} (T - y_T)/\sigma_T \\ \ell_Q/0.10 \\ g/0.10 \\ (T - T_c - 10\,\mathrm{K}\,\mathrm{e}^{\ell_Q - g})/\sigma_p \end{bmatrix},$$

$$\mathbf{J} = \begin{bmatrix} 1/\sigma_T & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 1/\sigma_p & -10\,\mathrm{e}^{\ell_Q - g}/\sigma_p & 10\,\mathrm{e}^{\ell_Q - g}/\sigma_p \end{bmatrix},$$

with the physics row of equation (3.19) carrying the tolerance $\sigma_p = 0.001\,\mathrm{K}$. At the prior means, $\mathbf{z}_0 = (T_c + 10\,\mathrm{K},\ 0,\ 0)^{\mathsf{T}}$, every residual is zero, the normal equations (3.25) give a zero step, and $\hat{\mathbf{z}} = \mathbf{z}_0$; the curvature term that Chapter 4 adds for nonlinear rows is a sum of residuals times second derivatives and vanishes with the residuals. There the Jacobian is

$$\mathbf{J} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 1000 & -10\,000 & 10\,000 \end{bmatrix},$$

whose last two columns carry the relation between load and conductance: both are zero in the reading row and equal and opposite in the physics row, so the direction $(0,\ 1,\ 1)^{\mathsf{T}}$, load and conductance raised together, gives $\mathbf{J}\,(0,\ 1,\ 1)^{\mathsf{T}} = (0,\ 10,\ 10,\ 0)^{\mathsf{T}}$ and is seen by the prior rows alone. Multiplying by $\mathbf{J}^{\mathsf{T}}$,

$$\mathbf{J}^{\mathsf{T}}\mathbf{J}\,(0,\ 1,\ 1)^{\mathsf{T}} = 100\,(0,\ 1,\ 1)^{\mathsf{T}},$$

so $(0,\ 1,\ 1)^{\mathsf{T}}$ is an eigenvector of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ with the eigenvalue $100 = 1/0.10^2$, the prior precision alone, and of $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ with the eigenvalue 0.0100, the prior variance: the reading has left

that direction untouched. Inverting $\mathbf{J}^\mathsf{T}\mathbf{J}$ numerically, equation (3.26) gives

$$\boldsymbol{\Sigma} = \begin{bmatrix} 0.039216 & 0.001961 & -0.001961 \\ 0.001961 & 0.005098 & 0.004902 \\ -0.001961 & 0.004902 & 0.005098 \end{bmatrix}.$$

For a Gaussian, the integral of equation (3.16) over $T$ leaves the block of $\boldsymbol{\Sigma}$ in the rows and columns of $\ell_Q$ and $g$, the lower right $2 \times 2$ block. Its precision is not the corresponding block of $\mathbf{J}^\mathsf{T}\mathbf{J}$ but, by the block-inversion formula of Section 3.5.3 applied to $\boldsymbol{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$, the Schur complement of the $T$ block, $\boldsymbol{\Lambda}_{qq} - \boldsymbol{\Lambda}_{qT}\boldsymbol{\Lambda}_{TT}^{-1}\boldsymbol{\Lambda}_{Tq}$ with $q$ the indices of $\ell_Q$ and $g$:

$$\boldsymbol{\Lambda}_{qq} - \boldsymbol{\Lambda}_{qT}\boldsymbol{\Lambda}_{TT}^{-1}\boldsymbol{\Lambda}_{Tq} = \begin{bmatrix} 2599.94 & -2499.94 \\ -2499.94 & 2599.94 \end{bmatrix} \longrightarrow \begin{bmatrix} 2600 & -2500 \\ -2500 & 2600 \end{bmatrix} \quad (\sigma_p \to 0),$$

against the block $\boldsymbol{\Lambda}_{qq}$ of $\mathbf{J}^\mathsf{T}\mathbf{J}$ itself, whose entries are $10^8 + 100$ on the diagonal and $-10^8$ off it. Each marginal standard deviation is $\sqrt{0.005098} = 0.0714$, down only from 0.100, and the correlation is $0.004902/0.005098 = 0.962$. The eigenvectors of the block are $(1,\ 1)^\mathsf{T}/\sqrt{2}$ and $(1,\ -1)^\mathsf{T}/\sqrt{2}$, with the eigenvalues $0.005098 \pm 0.004902$, that is 0.010000 and 0.000196: the posterior is as wide as the prior along the first and 7.1 times narrower along the second. The budget (3.34) with the gradients $\mathbf{g} = (0,\ 1,\ -1)^\mathsf{T}$ and $(0,\ 1,\ 1)^\mathsf{T}$ gives the variance of the difference, $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} = 0.000392$, a standard deviation of 0.0198, and that of the sum, 0.0200, a standard deviation of 0.1414, exactly its prior value $\sqrt{2} \times 0.10$. Repeated at $\sigma_p = 0.01\,\mathrm{K}$, the block entries print the same, 0.005098 and 0.004902, and the correlation prints 0.961 against 0.962 at 0.001 K; the limit is reached at 0.001 K.

*Remark* 3.2. Eliminating $T$ through the physics row turns the reading into the linearized row $(\ell_Q - g)/0.02$, since $\sigma_T/(T - T_c) = 0.02$, and the same block follows from a $3 \times 2$ Jacobian; along the difference alone, the prior precision 50 plus the reading's 2500 gives $1/2550 = 0.000392$.

Figure 3.7 draws the two distributions in the plane of the two entries. The horizontal axis is $\ell_Q$ and the vertical axis $g$; the dashed gray circle encloses 95 % of the prior, the green ellipse 95 % of the posterior, and the thin line is $\ell_Q - g = 0$, the value of the difference that the reading fixed. The posterior is a narrow ellipse lying along that line, its axes the two eigenvectors of the block: it is long in the direction the reading cannot see, $(1,\ 1)^\mathsf{T}$, where it reaches the prior circle because the reading added nothing there, and short in the direction it measured, $(1,\ -1)^\mathsf{T}$.

Read by its marginals, this posterior says that neither the load nor the conductance is well known, about 7 % each. Read by its shape, it says that their ratio, and with it the hall air at the present load, is known to 2 %. A prediction at a different, specified load $Q'$, $T' = T_c + Q'/G$, has a gradient with respect to $(T, \ell_Q, g)$ proportional to $(0,\ 0,\ 1)^\mathsf{T}$, so $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$ picks out $\Sigma_{gg}$ and the prediction inherits its 7 % spread; a prediction that scales the present load, such as the hall at twice the unknown $Q$, has a gradient proportional to $(0,\ 1,\ -1)^\mathsf{T}$, the short axis of the ellipse, and its temperature rise is known to the same 2 % relative as the present one; the absolute spread doubles. Which statement holds depends on the gradient of the question, and the covariance, not the list of marginal variances, carries the answer.

**Principle 3.4.** *Read the posterior's shape, not its marginals. Wide marginal spreads with a strong correlation mean that the evidence determines a combination of the entries; a quantity that depends on that combination is well known, and the entries are not unobservable merely because each alone is uncertain.*

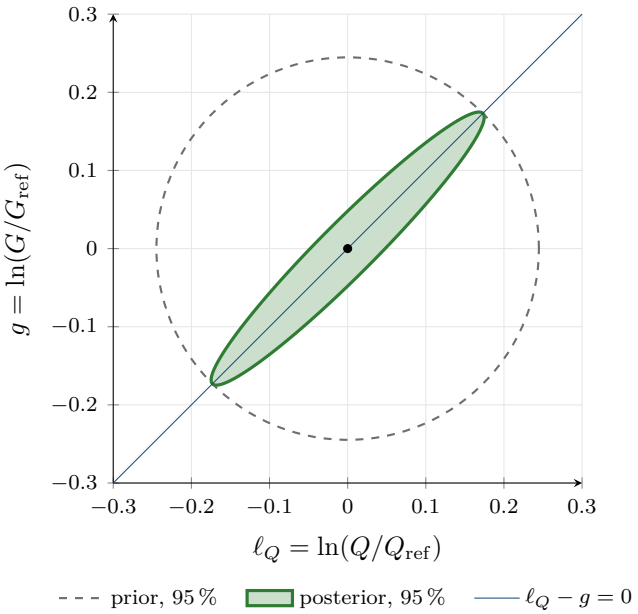



Figure 3.7: Prior and posterior of the log-load $\ell_Q$ and the log-conductance $g$ after one hall-air reading that fixes $\ell_Q - g$ within 0.02 (illustrative): the $(\ell_Q, g)$ block of the posterior of $\mathbf{z} = (T, \ell_Q, g)^{\mathsf{T}}$. Prior: independent, standard deviations 0.10. Posterior: standard deviations 0.0714, correlation 0.962; standard deviation of $\ell_Q - g$, 0.0198; of $\ell_Q + g$, 0.1414. Contours enclose 95 % ($\chi^2_{2,0.95} = 5.991$).

### 3.6.6 Mixtures over structure

A Gaussian approximation is most useful for continuous uncertainty near a well-identified posterior mode. Conditional probability is more general. If the plant may operate under either of two structures, $M \in \{M_1, M_2\}$ (two drawings of the plant), the law of total probability, applied over the models with $\mathbf{z}$ conditioned on the evidence throughout, gives

$$p(\mathbf{z} \mid \mathbf{y}) = \sum_j p(\mathbf{z}, M_j \mid \mathbf{y}) = \sum_j p(\mathbf{z} \mid \mathbf{y}, M_j)\ \Pr(M_j \mid \mathbf{y}), \tag{3.36}$$

where the second step is the multiplication rule. The weights follow from Bayes' theorem over the models, $\Pr(M_j \mid \mathbf{y}) = p(\mathbf{y} \mid M_j)\ \Pr(M_j) / \sum_i p(\mathbf{y} \mid M_i)\ \Pr(M_i)$, so that the evidence of each model, the normalizing constant of its own posterior, sets its share. The expression assumes that the relevant state $\mathbf{z}$ has a common interpretation across the models. Such a posterior need not be Gaussian, even when each model's posterior is.

A two-structure illustration on the hall air shows what is lost. Let the drawing $M_1$, the air path as designed, give $T \mid \mathbf{y}, M_1 \sim \mathcal{N}(26.9, 0.5^2)$, and a drawing $M_2$ with a degraded path give $\mathcal{N}(28.5, 0.5^2)$, with posterior model probabilities 0.7 and 0.3 (illustrative). By the law of total expectation the mixture mean is $0.7 \times 26.9 + 0.3 \times 28.5 = 27.38\,^{\circ}\mathrm{C}$; by the law of total variance, equation (3.32) with the model in place of the rate, its variance is the mean of the within-model variances plus the variance of the model means,

$$\mathrm{Var}(T \mid \mathbf{y}) = 0.5^2 + 0.7 \times 0.3 \times (28.5 - 26.9)^2 = 0.25 + 0.5376 = 0.7876\,\mathrm{K}^2,$$

a standard deviation of 0.887 K. The probability of exceeding 28 °C is the weighted sum of the two tails,

$$\Pr(T > 28\,^{\circ}\mathrm{C} \mid \mathbf{y}) = 0.7\,(1 - \Phi(2.2)) + 0.3\,(1 - \Phi(-1)) = 0.7 \times 0.0139 + 0.3 \times 0.8413 = 0.262.$$

The most probable drawing alone would report 0.014, and a single Gaussian with the mixture's mean and variance would report $1 - \Phi(0.699) = 0.242$ and hide the two separate modes, 3.2 standard deviations apart. The law of total probability combines the uncertainty over the structure with the uncertainty over the numerical state, and this combination becomes central in fault detection and isolation and in deciding which drawing of the plant the telemetry supports. The next chapter assembles these pieces into the complete workflow of probabilistic inference, from the plant graph and its likelihood to the posterior, its marginals, its expectations and the probabilities of functions of the state.

## 3.7 What to retain

**Definition 3.1** (Conditioning, marginalization and dependence)**.** A conditional probability restricts attention to the situations consistent with the evidence and renormalizes, $\Pr(A \mid B) = \Pr(A \cap B)/\Pr(B)$; Bayes' theorem reverses the conditioning through the prior and the evidence; marginalization sums or integrates over every unknown the question does not name; and the posterior covariance of a joint state, not its marginal variances, decides how uncertain a derived quantity is.

Four ideas carry forward. Control questions contain two kinds of uncertainty: aleatory uncertainty is the scatter of the next reading, load or trip count, and epistemic uncertainty is incomplete knowledge of the plant's gains, its conductance or a sensor's calibration; more windows reduce the second and leave the first, and the law of total variance separates them in a prediction. Conditional probability is how information is incorporated: Bayes' theorem updates a probability when an observation arrives, the base rate decides how far a trip moves it, and the fraction of trips that are real stiction is not the flag's probability of detection. Plant quantities are often dependent: a shared pump, a conservation relation or a common supply air creates relations among loops, units and unknown states, and conditional independence simplifies a calculation only where the model justifies it. A physics graph puts the shared physical causes in the state; the independence it still assumes is that of the errors, and a shared error, such as a calibration offset, is removed by carrying it as an entry too. The loop changes what that offset means: a controller holds its reading, so the offset neither averages away nor shows in the loop's error, and a tight trend is a statement about the loop. Marginalization answers a focused question without pretending that the other unknowns are known, whether they are nuisance quantities, competing conditions of a unit or competing drawings of the plant.

Together these establish the mechanism of inference: the prior $p(\mathbf{z})$ and the observation model $p(\mathbf{y} \mid \mathbf{z})$ combine into the posterior $p(\mathbf{z} \mid \mathbf{y})$, the update multiplying the prior by the likelihood and normalizing the product. For Gaussian rows the product is $\exp(-\frac{1}{2} r^{\mathsf{T}} r)$ over the whole state, physics rows included; the normal equations give its most probable state, $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ its covariance, exact when the rows are linear, and the conditional of the joint Gaussian of state and readings, through the Schur complement, gives the same answer. Relations between quantities are read in that space: a reading informs an entry it does not measure through a row of $\mathbf{J}$ that carries both, and a combination of entries is well known when it lies along a short eigenvector of $\boldsymbol{\Sigma}$, which $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ measures. Once the posterior is available, the engineer's question decides the next calculation: a marginal probability, an expectation, a prediction, a model comparison or a decision. Table 3.11 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 3.11: The terms of uncertainty and conditional probability and where each is defined.

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| Aleatory uncertainty | $\Pr(y_T > c \mid T) = 1 - \Phi((c - T)/\sigma_T)$ | – | readings vary with the hall known | 3.1.1 |
| Epistemic uncertainty | $p(g \mid \mathbf{y})$ | – | knowledge of a fixed quantity | 3.1.2 |
| Predictive exceedance | $\int \Pr(y_{\text{new}} > c \mid T)\, p(T \mid \mathbf{y})\, \mathrm{d}T$ | – | next reading, averaged over the hall | 3.1.3 |
| Conditional probability | $\Pr(A \cap B)/\Pr(B)$ | – | probability of $A$ among cases of $B$ | 3.2.2 |
| Multiplication rule | $\Pr(A \cap B) = \Pr(B \mid A)\Pr(A)$ | – | joint from conditional | 3.2.3 |
| Bayes' theorem | $\Pr(B \mid A)\Pr(A)/\Pr(B)$ | – | reverses the conditioning | 3.2.4 |
| Posterior odds | likelihood ratio × prior odds | – | Bayes without the evidence | 3.3.3 |
| Law of total probability | $\sum_s \Pr(H \mid s)\Pr(s)$ | – | marginal of an observation | 3.4.1 |
| Marginal likelihood | $\int p(\mathbf{y} \mid \lambda) p(\lambda)\, \mathrm{d}\lambda$ | – | evidence; normalizer | 3.4.2 |
| Clean-run probability | $\sum_j w_j\, \mathrm{e}^{-\lambda_{\text{ref}} \mathrm{e}^{\omega_j} \tau} = 0.96457$ | – | reported, no trip over the log-rate, equation (2.41) | 3.4.2 |
| Its exact value | $(b/(b+\tau))^a = 0.96863$ | – | same average over the gamma; gap 0.0041 | 3.4.2 |
| Nuisance marginal | $\int p(g, b_Q \mid \mathbf{y})\, \mathrm{d}b_Q$ | varies | integrate out, do not maximize | 3.4.3 |
| Conditional independence | $X_1 \perp X_2 \mid C$ | – | joint given $c$ factorizes | 3.4.6 |
| Common-cause joint | $\sum_c \Pr(F_A \mid c)\Pr(F_B \mid c)\Pr(c)$ | – | both loops trip with a shared cause | 3.4.5 |
| Correlated reading errors | $\mathbf{A}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{A}$, $\mathbf{R}$ full; or a shared bias entry | $\mathrm{K}^{-2}$ | the precision the readings truly add | 3.5.7 |
| Gaussian rows | $p(\mathbf{z} \mid \mathbf{y}) \propto \exp(-\frac{1}{2} r^{\mathsf{T}} r)$ | – | posterior of the whole state | 3.5.2 |
| Normal equations | $(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\boldsymbol{\Delta} = -\mathbf{J}^{\mathsf{T}} r$, $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ | varies | most probable state, covariance | 3.5.3 |
| Precisions add | $\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Sigma}_0^{-1} + \mathbf{J}_y^{\mathsf{T}}\mathbf{J}_y$ | varies | prior plus reading information | 3.5.3 |
| Gaussian conditional | $\boldsymbol{\Sigma}_{zz} - \boldsymbol{\Sigma}_{zy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yz}$ | varies | Schur complement | 3.5.3 |
| Exceedance probability | $1 - \Phi\big((T_{\max} - \mu_{\text{post}})/\sigma_{\text{post}}\big)$ | – | tail of the posterior | 3.5.5 |
| Law of total variance | $\mathbb{E}[\mathrm{Var}(K_\tau \mid \lambda)] + \mathrm{Var}(\mathbb{E}[K_\tau \mid \lambda])$ | – | process plus rate terms | 3.6.1 |
| Posterior covariance | $\boldsymbol{\Sigma}$, $\mathrm{corr}(z_1, z_2) = \Sigma_{12}/\sqrt{\Sigma_{11}\Sigma_{22}}$ | varies | joint spread of the state | 3.6.4 |
| Variance budget | $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ | varies | sensitivity$^2$ × variance, plus cross terms | 3.6.4 |
| Mixture posterior | $\sum_j p(\mathbf{z} \mid \mathbf{y}, M_j)\Pr(M_j \mid \mathbf{y})$ | varies | uncertainty over structure | 3.6.6 |

## 3.8 Exercises

**Exercise 3.1** (Formulation)**.** (a) In 2000 comparable loop-runs of a trip log the flag fired in 100, real stiction occurred in 60, and both in 25. Compute $\Pr(H)$, $\Pr(E)$, $\Pr(H \cap E)$,

$\Pr(E \mid H)$ and $\Pr(H \mid E)$, name what the last two mean for the flag, and state whether trip and stiction are independent. (0.05, 0.03, 0.0125, $25/100 = 0.25$, the fraction of trips that are real stiction, and $25/60 = 0.417$, the probability of detection; not independent, since $0.0125 \neq 0.05 \times 0.03 = 0.0015$.) (b) Two loops on one chilled-water pump have $\Pr(F_A) = 0.05$, $\Pr(F_B) = 0.04$ for a trip in a run and $\Pr(F_A \cap F_B) = 0.015$. Compute the joint probability under an incorrect independence assumption, the probability that at least one loop stays in automatic, and $\Pr(F_B \mid F_A)$, and state what it means for the night. ($0.05 \times 0.04 = 0.002$; $1 - 0.015 = 0.985$, against 0.998 under independence; $0.015/0.05 = 0.30$; both loops trip together 7.5 times as often as independence predicts, so a night with both loops needing the engineer is far more likely than the product suggests.)

**Exercise 3.2** (Formulation)**.** A hall satisfies $T = T_c + Q/G$ with $T_c$ known, $G > 0$, and both $Q$ and $G$ uncertain. Starting from $\Pr(T > T_{\max} \mid \mathbf{y})$, write an equivalent event in $Q$ and $G$ and express the probability as an integral over their joint posterior, without assuming independence. (Since $G > 0$, $T_c + Q/G > T_{\max} \iff Q > G(T_{\max} - T_c)$, so $\Pr(T > T_{\max} \mid \mathbf{y}) = \iint \mathbf{1}[Q > G(T_{\max} - T_c)]\, p(Q, G \mid \mathbf{y})\, \mathrm{d}Q\, \mathrm{d}G$; the joint posterior keeps the correlation between $Q$ and $G$ that a reading of $T$ creates.)

**Exercise 3.3** (Derivation)**.** (a) A fan-vibration trip monitors a cooling unit with $\Pr(F) = 0.02$ per run, $\Pr(H \mid F) = 0.95$ and $\Pr(H \mid \neg F) = 0.05$. Derive the probability that the fan is faulty when the trip activates, and explain why it is not 95 %. ($\Pr(H) = 0.95 \times 0.02 + 0.05 \times 0.98 = 0.068$; $\Pr(F \mid H) = 0.019/0.068 = 0.279$; 0.95 is $\Pr(H \mid F)$, not $\Pr(F \mid H)$, and the low base rate makes most trips false.) (b) With prior $\Pr(F) = 0.10$ that the unit is degraded, the valve-position trip $A$ has $\Pr(H_A \mid F) = 0.9$, $\Pr(H_A \mid \neg F) = 0.1$, and the hall-temperature trip $B$ has $\Pr(H_B \mid F) = 0.8$, $\Pr(H_B \mid \neg F) = 0.2$; the trips are conditionally independent given the unit's state. Derive $\Pr(F \mid H_A, H_B)$ and explain why the assumption is needed. ($\Pr(H_A, H_B \mid F) = 0.9 \times 0.8 = 0.72$ and $\Pr(H_A, H_B \mid \neg F) = 0.1 \times 0.2 = 0.02$, so $\Pr(F \mid H_A, H_B) = 0.072/(0.072 + 0.018) = 0.80$; sequentially, trip $A$ gives 0.5 and trip $B$ then 0.8; the likelihoods multiply only under conditional independence given each state of the unit, and a control fault that raises both trips whatever the unit's state would invalidate it.)

**Exercise 3.4** (Derivation)**.** The count of future trips satisfies $K_\tau \mid \lambda \sim \text{Poisson}(\lambda\tau)$ with $\lambda \mid \mathbf{y} \sim \text{Gamma}(3, 1000\,\text{h})$ (shape and rate) and $\tau = 200\,\text{h}$, about seventeen runs. Derive the expected count, the predictive variance and its two contributions, and state what happens to each when a much longer trip log supporting a similar rate arrives. (The expected count is $\mathbb{E}[K_\tau \mid \mathbf{y}] = 200 \times 0.003 = 0.6$; by equation (3.33), the predictive variance is $0.6 + 200^2 \times 3/1000^2 = 0.6 + 0.12 = 0.72$; process 0.6, rate 0.12, a share $200/1200 = 0.167$. With $a$ and $b$ growing in proportion, the rate term $\tau^2(a/b)/b$ tends to zero while the process term stays near 0.6.)

**Exercise 3.5** (Application)**.** A loop operates in mild weather with probability 0.90 per run and a trip probability of 0.01, and in hot weather, with the valve near its stop, with probability 0.10 and a trip probability of 0.15. Derive the unconditional trip probability of a run, and the probability that the weather was hot given that a trip occurred. ($\Pr(E) = 0.90 \times 0.01 + 0.10 \times 0.15 = 0.024$; $\Pr(\text{hot} \mid E) = 0.015/0.024 = 0.625$.)

**Exercise 3.6** (Application)**.** (a) A hall air temperature has prior $\mathcal{N}(25, 1.5^2)$, and a sensor with $\sigma_T = 0.5\,\text{K}$ reads $y_T = 26°\text{C}$. For the state, the one number $T$, write the two residuals and their slopes $j_k$, the information $\Lambda$, take the Newton step from the prior mean, and compute $\sigma^2_{\text{post}}$ and $\Pr(T > 27°\text{C} \mid y_T)$. ($r_1 = (T - 25)/1.5$ and $r_2 = (T - 26)/0.5$, $j_1 = 0.6667$ and $j_2 = 2\,\text{K}^{-1}$, $\Lambda = 0.4444 + 4 = 4.4444\,\text{K}^{-2}$; at $T = 25$, $r_1 = 0$ and $r_2 = -2$, so $\mathcal{L}' = j_1 r_1 + j_2 r_2 = -4$, $\Delta T = 4/4.4444 = 0.9$ and $\hat{T} = 25.90°\text{C}$; $\sigma^2_{\text{post}} = 1/4.4444 = 0.225\,\text{K}^2$, $\sigma_{\text{post}} = 0.4743\,\text{K}$; $1 - \Phi(1.10/0.4743) = 1 - \Phi(2.319) = 0.0102$.) (b) For the two-entry hall of equation (1.10), whose Jacobian at the forward state (23°C, 0) has the rows (100, 500), (2, 0) and (0, 3.333),

compute the three outer products $\mathbf{J}_k^{\mathsf{T}}\mathbf{J}_k$ and their sum, and state which row creates the off-diagonal entry and what the rank of its term is. (Physics $\begin{bmatrix}10\,000 & 50\,000\\ 50\,000 & 250\,000\end{bmatrix}$, reading diag(4, 0), prior diag(0, 11.11); sum $\begin{bmatrix}10\,004 & 50\,000\\ 50\,000 & 250\,011\end{bmatrix}$; only the physics row has entries in both columns, and its term has rank one, determinant $10^4 \times 2.5 \times 10^5 - (5 \times 10^4)^2 = 0$.)

# Chapter 4

# The Anatomy of Probabilistic Inference

## 4.1 From a plant graph to a release probability

The commissioning engineer of the equipment hall has the record of its acceptance test open: the discharge loop holding the supply air at its set point of 18°C, a test load of 30 kW, and one hall sensor reading 23.2°C. The hall is to be released for 60 kW against an acceptance limit of 28°C. The sheet has room for one word, release or refuse, and the two entries already on it point opposite ways. The design sheet's handbook value, a resistance of 0.160 K/kW for the hall's air path, puts the hall at $18 + 60 \times 0.160 = 27.6$°C at full load: release. The test's rise of 5.2 K above the supply air, doubled with the load, puts it at 28.4°C: refuse. These are the commissioning record's own entries, not the values and the limit of Chapter 1.

Neither entry says how sure the engineer can be: the design sheet ignores the test, and the doubled rise ignores the design sheet and the error of the sensor. What neither settles is what one test, the sheet and the physics together say about a load that has never been run, and how far the set point would have to move if the hall is to be released with the load held. The sibling volume *Operations and Control by Derivation* asks the same hall for a probability and stops there; for a control engineer the set point is the handle, so the answer has to end on where to put it. The question of this chapter is

*how a posterior distribution over the state of a plant graph is obtained from physical relations, readings and prior information, and how that posterior is turned into the probability, and the set point, that answer a control engineer's question.*

Chapter 1 built the plant graph, Chapter 2 classified the questions and Chapter 3 derived the rules one at a time; here they are used together for the first time, on a state of three entries of which the reading touches only one.

**Jargon: state estimation, maximum a posteriori (MAP) estimate, Hessian, Laplace approximation, delta method, change of variables, credible interval, prior and posterior predictive distributions, marginal likelihood, Bayes factor, rank-one update, reduced state, scenario node, back-off**

*State estimation* is the fit of the state of a plant to its readings and priors; its results are the MAP state and the posterior covariance defined next. The *maximum a posteriori (MAP) estimate* is the value of the unknowns at which the posterior density is highest; it is found by minimizing the negative logarithm of the posterior, the *objective*. The *Hessian* is the matrix of second derivatives of the objective; at the MAP estimate it measures the curvature of the objective, and for rows linear in the state and divided by their standard deviations it is $\mathbf{J}^\mathsf{T}\mathbf{J}$, the information matrix. The *Laplace approximation* replaces the posterior by the Gaussian whose mean is the MAP estimate and whose covariance is the inverse of the Hessian there. The *delta method* approximates the variance of a function of the unknowns by linearizing the function at the estimate. A *change of variables* carries a density from one coordinate to another, multiplying it by the derivative of the old coordinate with respect to the new so that each interval keeps its probability. A *credible interval* contains the unknown with a stated posterior probability. The *prior predictive* and *posterior predictive* distributions are the distributions of a reading not yet seen, averaged over the prior and over the posterior respectively. The *marginal likelihood*, or *evidence*, of a model is the probability density of the observed readings averaged over the prior; the *Bayes factor* is the ratio of the evidences of two models. A *rank-one update* adds one new reading to an existing Gaussian posterior by a closed-form correction of its mean and covariance, without solving the problem again. The *reduced state* is the vector of free entries that remain when the physical rows are taken as exact; every other entry of the state is then a function of it. A *scenario node* is the plant re-solved with an input pinned to a value not yet applied. The *back-off* is the distance a set point stands inside its limit because the controlled quantity is uncertain.

**Key idea**

A release probability is computed in two operations. Inference, the state estimation of the plant, turns the plant graph, the readings and the priors into a posterior over the state: its mean is the MAP state and its covariance is the inverse curvature of the objective. Query evaluation turns that posterior into the engineer's answer: an entry of the state and its variance, or a derived quantity $f(\mathbf{z})$ with variance $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$, and then a normal tail at the limit.

### 4.1.1 The complete inference workflow

The method fixed in Chapter 2 (Figure 2.6) is a progression: the plant graph (step 0), then the questions written over its state (step 1), the graph extended for them with the evidence attached (step 2), the posterior of the state that the rows give, derived and computed once (step 3), the answers read from that posterior (step 4), their checks (step 5), the report with its bands (step 6), and the next reading when a band is too wide (step 7). Figure 4.1 draws the progression. The wide box at the top is the plant graph, built once; the seven boxes beneath are one pass that carries every question, read left to right along the top row and right to left along the bottom row; the second line of each box names what the step produces. The dashed arrow from step 7 back to step 3 is the loop taken when a new reading is requested: its row is added and the posterior is rebuilt.

The distinction between steps 3 and 4 carries the method, and it is the distinction between the two operations of every inference. Step 3 is *inference*: the rows of step 2 are combined into one posterior of the state, its MAP state, information and covariance, derived and then computed for the plant; it does not depend on which question is asked. Step 4 is *query evaluation*:

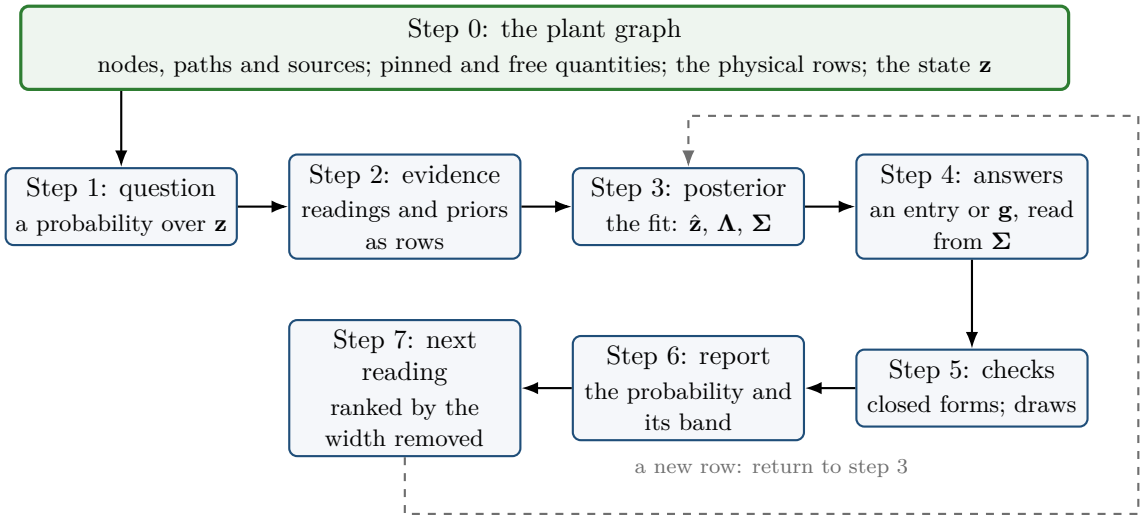



Figure 4.1: The inference workflow. Step 0 is built once; steps 1 to 7 are one pass, carrying every question of the plant. The second line of each box names what the step produces.

each question is reduced, from the engineer's statement through its general and specialized forms, to quantities read from that posterior, in this chapter the posterior mean and variance of one state entry, the full-load temperature, and the standard normal distribution function evaluated at the limit. Keeping them apart is what lets one posterior serve every question, and a new reading change step 3 alone.

### 4.1.2 The posterior and the questions put to it

The calculation divides into two operations of different kinds:

$$\underbrace{(\text{graph},\ \mathbf{y}) \longrightarrow p(\mathbf{z} \mid \mathbf{y})}_{\text{inference}}, \qquad \underbrace{p(\mathbf{z} \mid \mathbf{y}) \longrightarrow \text{probability or decision}}_{\text{query evaluation}}, \tag{4.1}$$

the second applied to the result of the first. Inference is done once for a given graph and a given body of evidence. Under a Gaussian posterior, exact or approximate, its result is the triple of the MAP state $\hat{\mathbf{z}}$, the posterior covariance $\boldsymbol{\Sigma}$, and the Jacobian of the model at $\hat{\mathbf{z}}$, from which the gradient $\mathbf{g}$ of any derived quantity is formed. Query evaluation is then a formula on the triple. An output that is itself an entry $z_i$ of the state, such as a temperature at a node, has mean $\hat{z}_i$ and variance $\Sigma_{ii}$, read straight off the posterior. An output that is a function $f(\mathbf{z})$ of several entries needs its gradient and the quadratic form $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$. A probability that the output stays below a limit is a normal tail on either pair.

Two kinds of further work follow from this division and must be kept apart. A new question asked of the same evidence needs no new inference: a new function of the state is pushed through the existing $\boldsymbol{\Sigma}$. New evidence changes the posterior itself: either the inference is repeated with the new row, or the existing Gaussian is conditioned on the new row by a rank-one update, which Section 4.6.4 derives. The posterior represents what is known about the plant under the model; the queries represent what the engineer wants to determine.

### 4.1.3 The inference in the terms of graphical models

The operations of this book are those of inference in a probabilistic graphical model, under the names a control engineer uses. The plant graph with its rows is a factor graph (Section 1.1.3): each row is a factor over the entries it touches, and the posterior $\mathrm{e}^{-\mathcal{L}(\mathbf{z})}$ is the normalized product of the factors. Table 4.1 sets each operation beside its name in the literature on graphical

models. Its first column is the operation as this book writes it, the second the standard name, and the third the chapters that derive it.

Table 4.1: The operations of the book and their names in the literature on probabilistic graphical models.

| Operation in this book | Name in graphical models | Derived in |
|---|---|---|
| a row: prior, physics or reading | a factor of a factor graph | Chapters 1, 3 |
| the posterior $e^{-\mathcal{L}(\mathbf{z})}$ | the normalized product of the factors | Chapter 3 |
| the normal equations, $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$, $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$ | exact inference in a linear-Gaussian factor graph, in information form | Chapter 3; this chapter |
| the gamma updates of a trip rate | exact inference with conjugate factors | Chapters 2, 3 |
| Gauss–Newton and the Laplace covariance | the Gaussian approximation at the mode of a nonlinear factor graph (factor-graph optimization) | Chapter 1; this chapter |
| the Schur complement; the marginal of a joint Gaussian | variable elimination of the entries not asked about | Chapter 3 |
| the ranking of fault hypotheses; total probability | exact inference over discrete hypotheses by summation | Chapter 27 |
| the evidence $p(\mathbf{y} \mid M)$ | the normalizing constant (partition function) of the product | Chapter 2 |
| the Laplace evidence against draws | importance sampling, with the Laplace normal as the proposal | not derived in this book |

The general algorithms of that literature are not needed for the plants of this book, and the table shows why: every posterior here is solved jointly, by factoring the information matrix of the whole state, which is exact for linear-Gaussian rows and costs little at a few dozen entries. Message passing, belief propagation (the sum-product algorithm) and the junction tree compute the same marginals by local operations on the graph, and they earn their place on graphs too large to factor jointly or with discrete entries of many levels; on a Gaussian graph without loops, belief propagation returns exactly the Schur complements of the joint factorization. Factoring a large sparse information matrix is itself variable elimination, and its cost is set by the order in which the entries are eliminated: eliminated in time order, a state indexed by time, such as the controller steps of a loop trend, gives the Kalman filter and smoother (Chapter 9). Markov chain Monte Carlo (Gibbs sampling and the Metropolis algorithm) draws from a posterior that the Laplace normal cannot approximate and reweighting cannot correct: one with several modes, or the funnel of a hierarchical model, in which the spread of a population of rates is itself unknown; a later volume of *Reliability Engineering by Derivation* meets it in a fleet of pumps. Variational inference is not used in this book.

These algorithms differ in what they give up. Variable elimination, the junction tree, belief propagation on a graph without loops and the Kalman smoother on linear-Gaussian rows are exact: they compute the posterior the joint factorization computes, in another order. The Laplace approximation, loopy belief propagation, the linearized Kalman filters and variational inference are approximate, and their error is a bias that more computation does not remove; loopy belief propagation and variational inference in particular tend to report variances too small, which for a tail probability against a limit is the error that matters most. Importance sampling and Markov chain Monte Carlo are exact in the limit of many draws, and their error is random and can be estimated, although a chain that never visits a second mode can look converged. The two kinds of error check each other, which is why the draws serve as the other route against the Laplace answer in step 5. The route also decides which objects come cheaply. The joint factorization gives $\hat{\mathbf{z}}$, $\mathbf{\Sigma}$, the gradient and the preposterior covariance $\mathbf{\Sigma}_a$ of a rank-one

update in one pass, and the evidence in closed form under the Laplace approximation; draws give the posterior as a sample, from which a variance is an estimate, a preposterior needs the candidate's outcomes simulated and the posterior refitted or reweighted, and the evidence needs a further computation, since the draws do not carry the normalizing constant. The method is unchanged by the choice: the plant graph, the questions, the rows and the objective $\mathcal{L}(\mathbf{z})$ are the same for every algorithm, the general forms of Chapter 2 are expectations over the posterior whatever computes it, and only the computations of steps 3 and 4, and the checks of step 5 that test them, depend on the route.

### 4.1.4 The plant graph under two operating conditions (step 0)

The plant is an equipment hall with a heat-generating load and a single effective heat-removal path to the supply air, which the cooling unit's discharge controller holds at $T_c = 18°\mathrm{C}$. Two operating conditions are considered: a test condition with a heat load of $Q_1 = 30\,\mathrm{kW}$, and the requested full load $Q_2 = 60\,\mathrm{kW}$. At steady state the heat supplied to the hall leaves it through the path, and energy conservation at the hall node reads

$$Q = G\,(T - T_c), \tag{4.2}$$

where $T$ is the hall air temperature in °C, $Q$ the heat load in kW, and $G$ the effective thermal conductance of the air path in kW/K. The supply air is pinned: the controller holds it at its set point $r$, here $r = T_c = 18°\mathrm{C}$, and this part of the book treats that controller as ideal, which is what makes $T_c$ a boundary of the graph rather than an unknown, and it fails if the unit runs out of capacity and the discharge drifts up, which Section 4.6.6 takes up. The balance is taken at steady state through one effective path: this reduces the cooling of the hall to a single unknown, and it fails in the minutes after a load change, while heat stored in the hall and its equipment is still moving, which Section 4.6.6 and Exercise 4.6 take up. It is convenient to work with the thermal resistance of the path,

$$\rho = \frac{1}{G}, \tag{4.3}$$

in K/kW, with which equation (4.2) becomes

$$T = T_c + Q\rho. \tag{4.4}$$

Carrying the resistance in the state, rather than the conductance, is a choice: the temperature is then a linear function of the state, which makes every row of the example linear, the posterior exactly Gaussian and every result below exact rather than approximate, and it fails only where a Gaussian on the resistance puts real probability below zero, which Section 4.1.7 measures for the handbook prior and Section 4.6.6 discusses. Chapter 1 carries the conductance in the state as the log-conductance $g = \ln(G/G_{\mathrm{ref}})$, which keeps it positive but makes the temperature nonlinear in the state.

At the two operating conditions equation (4.4) gives

$$T_1 = 18 + 30\rho, \tag{4.5}$$

$$T_2 = 18 + 60\rho, \tag{4.6}$$

with $T_1$ the hall temperature during the test and $T_2$ the temperature at full load, both in °C with $\rho$ in K/kW. Figure 4.2 draws the graph. Each row is one operating condition: the box on the left is the heat source, the green box in the middle the hall node carrying a free temperature, and the gray box on the right the supply-air node, pinned at 18°C. The solid arrows are the physical paths, the heat entering the hall and the heat leaving it through the resistance. The green node between the rows is the resistance $\rho$, free and shared by both paths,

as the dashed links show. The dashed arrows leave the graph for the evidence: the reading of the test temperature and the handbook prior on $\rho$, attached in Section 4.1.7. The full-load node carries no reading; it is the scenario node of a what-if (contingency) analysis that the question of Section 4.1.6 requires, drawn here with the rest of the graph so that the figure shows the graph as extended for that question.

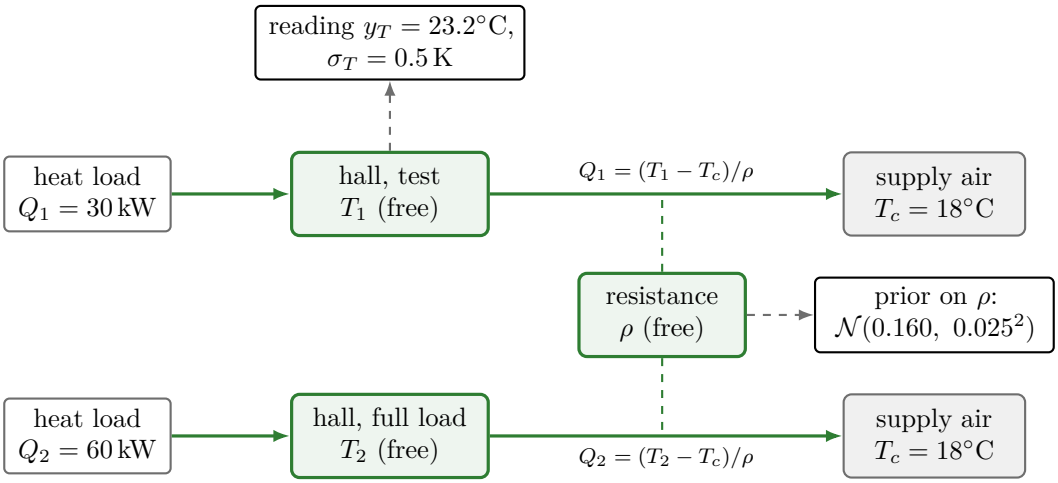



Figure 4.2: The plant graph of the equipment hall under its two operating conditions. Free: the resistance $\rho$ and the temperatures $T_1$ and $T_2$ (green). Pinned: the supply-air temperature $T_c$, held by the cooling unit's discharge controller, and the loads $Q_1$, $Q_2$ (gray and white). Observed: $T_1$, by one reading (dashed arrow, top); $\rho$ carries the handbook prior (dashed arrow, right). The full-load node $T_2$ is the scenario node of the question and has no reading; it is linked to the test only through the shared resistance.

The resistance is taken to be the same at 60 kW as at 30 kW, one free entry shared by both rows: this is what lets a reading taken at the test condition carry information about the full-load condition, although the hall has never been operated at full load, and it fails if the path changes with load, for example when the unit's fans or its coil change state, which one test reading cannot detect (Section 4.6.6) and a reading at full load would test (Section 4.6.3). The link between the two rows of Figure 4.2 is not a device of the calculation; it is the physical statement that both conditions are the same hall with the same heat-removal path, and the prediction at full load rests on it.

### 4.1.5 The state vector and the physical rows

The free quantities of the graph form the state vector,

$$\mathbf{z} = \begin{bmatrix} \rho \\ T_1 \\ T_2 \end{bmatrix}, \tag{4.7}$$

with entries in K/kW, °C and °C. The physical relations are written as rows that vanish when they hold:

$$r_1(\mathbf{z}) = T_1 - 18 - 30\rho = 0, \tag{4.8}$$
$$r_2(\mathbf{z}) = T_2 - 18 - 60\rho = 0. \tag{4.9}$$

These are the conservation rows of the graph, one for each hall node. The balance is taken to hold at both conditions up to a small model tolerance $\sigma_p$, a standard deviation that represents model discrepancy, and the answers are reported in the limit $\sigma_p \to 0$ of exact balances: the

finite tolerance keeps every matrix of the calculation invertible while the reported numbers are those of an exact balance, and the limit understates the spread if the real discrepancy is finite, which Section 4.4.5 measures over three tolerances and Section 4.6.6 discusses. Section 4.2.3 puts the rows, with their tolerance, into one residual vector with the reading and the prior.

The three entries of the state are therefore not three independent unknowns. The two physical rows leave one direction of the state free, a change of $\rho$ with both temperatures following it; in the limit of exact rows the state has three entries but one degree of freedom, and this fact governs the shape of the posterior covariance derived in Section 4.2.5 and the reduced state of Section 4.2.8.

### 4.1.6 The control engineer's questions (step 1)

The release raises the heat load from 30 kW to 60 kW. The maximum acceptable hall temperature, the acceptance limit of the commissioning test, is $T_{\max} = 28°\mathrm{C}$. In the engineer's words the question is: given the hall sensor's reading and what is known of the air path, what is the probability that the hall stays below 28°C at full load? As a probability over the state, with the unknown on the left of the bar and everything treated as known on the right,

$$\Pr(T_2 < T_{\max} \mid \mathbf{y}, M), \tag{4.10}$$

where $T_2$ is the full-load entry of the state, $\mathbf{y}$ the evidence (the test reading and the handbook prior), and $M$ the assumed steady-state thermal model, equation (4.4) with the rows (4.8)–(4.9). The model is suppressed from the notation below when it is not in question. This is a question of the category *prediction under a threshold*: a tail probability of one marginal of the posterior.

Three estimation questions are asked of the same evidence, one for each entry of the state: the thermal resistance of the air path, the hall temperature at the test condition, and the hall temperature at full load. A fifth asks what the handle can do about the answer: with the set point pinned at a value $r$ not yet applied, the full-load temperature is $T_2^*(r) = r + Q_2\rho$, the scenario node of the question, and the set point is to be placed so that this temperature stays below the limit with probability 0.95. Table 4.2 lists the five questions: the first column is the question in the control engineer's words; the second the general form that every question of its category takes, written for a quantity $f(\mathbf{z})$ derived from the state and a threshold $f^*$ on it; the third its mathematical form for the hall; the fourth its category; and the last the subsections where it is answered. Which further reading would narrow the answers most asks nothing new of the present evidence: it is the next reading for these five questions, answered in Section 4.6.

Because the full-load temperature is carried as an entry of the state, the threshold question needs exactly one marginal of the posterior: the mean $\hat{z}_3$ and the variance $\Sigma_{33}$ of the third entry: the question is written over the state, and the position of $T_2$ in the state shows which part of the posterior it needs. Each estimation question decomposes in the same way, into one entry of $\hat{\mathbf{z}}$ and one diagonal entry of $\boldsymbol{\Sigma}$.

### 4.1.7 The evidence as rows (step 2)

**The temperature reading.** During the 30 kW test the hall sensor reports $y_T = 23.2°\mathrm{C}$, with a measurement error of standard deviation $\sigma_T = 0.5\,\mathrm{K}$. The observation model writes the random reading $Y$ as the true test temperature, the second entry of the state, plus an error,

$$Y = T_1 + \epsilon = \mathbf{h}^{\mathsf{T}}\mathbf{z} + \epsilon, \qquad \mathbf{h} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \epsilon \sim \mathcal{N}(0, \sigma_T^2), \tag{4.11}$$

where $\mathbf{h}$ selects the entry the sensor measures. The error is taken as Gaussian, unbiased, of known spread and independent of the state: this makes the reading one linear row of fixed

Table 4.2: The questions asked of the hall's test reading, the general form of their category, their mathematical form for the hall, their category and where each is answered. $\mathbf{y}$ is the test reading and the handbook prior; $T_{\max} = 28°\mathrm{C}$.

| Control engineer's question | General form | Mathematical form | Category | Answered in |
|---|---|---|---|---|
| What is the thermal resistance of the air path? | moments, (2.10) | $p(\rho \mid \mathbf{y})$: $\hat{z}_1$, $\sqrt{\Sigma_{11}}$ | estimation | 4.2.10, 4.3.6 |
| What was the hall temperature during the test? | moments, (2.10) | $p(T_1 \mid \mathbf{y})$: $\hat{z}_2$, $\sqrt{\Sigma_{22}}$ | estimation | 4.3.6 |
| What will the hall temperature be at full load? | moments, quantiles, (2.10), (2.11) | $p(T_2 \mid \mathbf{y})$: $\hat{z}_3$, $\sqrt{\Sigma_{33}}$, credible interval | estimation | 4.3.1, 4.3.4, 4.3.6 |
| Can the hall be released for 60 kW within 28°C? | threshold, (2.12) | $\Pr(T_2 < T_{\max} \mid \mathbf{y})$, with $f = -T_2$ | prediction under a threshold | 4.3.5, 4.3.7 |
| How far must the set point move to hold it at 0.95? | threshold at a scenario node, (2.12) | $r_{95}$ from $\Pr(T_2^*(r) < T_{\max} \mid \mathbf{y}) = 0.95$ | prediction under a threshold | 4.3.8 |

weight $1/\sigma_T^2$, and it fails if the sensor is biased, which would shift every inferred quantity; the prior predictive check of Section 4.4.2 would flag a large bias, and the comparison of models of Section 4.2.2 is the test built for it. The reading lands on $T_1$ only; it is the physical row of the test hall, not a substitution, that will carry it to $\rho$.

**Prior information.** The design sheet gives a preliminary value of the resistance, the handbook figure for an air path of this kind, represented by the Gaussian prior

$$\rho \sim \mathcal{N}(\mu_0, \sigma_0^2), \qquad \mu_0 = 0.160\,\mathrm{K/kW}, \quad \sigma_0 = 0.025\,\mathrm{K/kW}. \tag{4.12}$$

The prior mean corresponds to a conductance $G_0 = 1/0.160 = 6.25\,\mathrm{kW/K}$. Taking the handbook value as the prior for this hall is a choice: it is what places the full-load temperature at 27.6°C before any reading and what keeps the estimate from following the one reading alone, and it misleads if the handbook describes air paths unlike this one, a conflict that the prior predictive check of Section 4.4.2 would show as a reading far from 22.8°C. A Gaussian prior on a resistance assigns some probability to negative, nonphysical values; here it is $\Phi(-0.160/0.025) = \Phi(-6.4) \approx 8 \times 10^{-11}$, negligible, and the Gaussian is kept for its analytical simplicity. A production model enforces positivity through a logarithmic parameterization, such as the log-conductance $g$ of Chapter 1, or a prior with positive support.

**The rows and the entries they act on.** The physics and the evidence put four rows on the state: the physical rows of the test hall and of the full-load hall, equations (4.8)–(4.9), the reading row on $T_1$, and the prior row on $\rho$. Which row acts on which entry is a property of the graph, and it fixes where the Jacobian of the rows can be nonzero, marked $*$:

$$\mathbf{J} = \begin{bmatrix} * & * & 0 \\ * & 0 & * \\ 0 & * & 0 \\ * & 0 & 0 \end{bmatrix} \begin{matrix} \text{physics, test hall} \\ \text{physics, full-load hall} \\ \text{reading} \\ \text{prior} \end{matrix}, \qquad \text{columns } \rho, T_1, T_2. \tag{4.13}$$

A line of the pattern lists the entries one row sees, and a column lists the rows that see one entry. The third column has a single mark: the full-load temperature is seen by its own physical

row and by nothing else. The second column shows that the reading and the test hall's balance are the two rows that see $T_1$, and the first that the two balances and the prior see $\rho$. No row joins the reading to $T_2$; the two physical rows, through the column $\rho$ they share, are the only path from the one to the other.

**Knowns and unknowns.** Table 4.3 lists every quantity of the problem. The first column is the symbol, the second its value, and the third its role in the graph: pinned boundary condition, evidence, or free entry of the state. The graph and the probabilistic model are now specified completely enough for the inference to be carried out.

Table 4.3: Knowns and unknowns of the equipment-hall example.

| Quantity | Value | Role |
|---|---|---|
| $T_c = r$ [°C] | 18 | pinned: supply air, held by the discharge controller at its set point $r$ |
| $Q_1$ [kW] | 30 | pinned: test load |
| $Q_2$ [kW] | 60 | pinned: requested load |
| $y_T$ [°C] | 23.2 | evidence: the test reading |
| $\sigma_T$ [K] | 0.5 | evidence: standard deviation of the reading error |
| $\mu_0$ [K/kW] | 0.160 | evidence: prior mean of the resistance |
| $\sigma_0$ [K/kW] | 0.025 | evidence: prior standard deviation of the resistance |
| $\sigma_p$ [K] | 0.01, then $\to 0$ | model: tolerance of the two physical rows |
| $\rho$ [K/kW] | unknown | free: thermal resistance, the first entry of the state |
| $T_1$, $T_2$ [°C] | unknown | free: test and full-load temperatures, tied to $\rho$ by the physical rows |

*Remark* 4.1. Four spreads appear in what follows and must not be confused. The reading error has standard deviation $\sigma_T$ (no index), a property of the sensor. The prior on the resistance has standard deviation $\sigma_0$. The physical rows have the model tolerance $\sigma_p$, a property of the steady-state model. The posterior standard deviation of the $i$-th entry of the state is $\sqrt{\Sigma_{ii}}$, the square root of a diagonal entry of the posterior covariance, written $\mathrm{sd}(T_2 \mid \mathbf{y})$ when the entry is named.

## 4.2 The posterior of the state (step 3)

Every question of Table 4.2 is a property of one density, the posterior of the state given the reading and the handbook prior. Each row of Section 4.1.7 contributes one factor to it, and the MAP state, the information matrix and the covariance are properties of that product. This section derives the posterior and computes it for the hall; Section 4.3 reads every answer from it.

### 4.2.1 The likelihood of the reading and the physical rows

The likelihood states how probable the observed reading would be under a proposed state. By equation (4.11), for a fixed state the reading is the test temperature $T_1$ plus a Gaussian error, so $Y \mid \mathbf{z} \sim \mathcal{N}(T_1,\ \sigma_T^2)$, and its density at the observed value is

$$p(y_T \mid \mathbf{z}) = \frac{1}{\sqrt{2\pi\sigma_T^2}} \exp\left[-\frac{(T_1 - y_T)^2}{2\sigma_T^2}\right]. \tag{4.14}$$

For the reading $y_T = 23.2°\mathrm{C}$ and $\sigma_T = 0.5\,\mathrm{K}$, so that $2\sigma_T^2 = 0.5\,\mathrm{K}^2$,

$$p(y_T \mid \mathbf{z}) \propto \exp\left[-\frac{(T_1 - 23.2)^2}{0.5}\right]. \tag{4.15}$$

The likelihood is a function of the state with the reading held fixed; it is not a density in $\mathbf{z}$ and need not integrate to one over it. It depends on the first temperature only and is flat in $\rho$ and $T_2$: the sensor measures a temperature and says nothing directly about the resistance or about the full-load hall.

What connects the three entries is the pair of physical rows. A steady-state balance that holds up to a model tolerance $\sigma_p$ (a standard deviation in kelvin, the amount by which the balance of equation (4.4) is allowed to fail) is the density

$$p(T_1 \mid \rho) = \mathcal{N}(T_1;\ 18 + 30\rho,\ \sigma_p^2), \qquad p(T_2 \mid \rho) = \mathcal{N}(T_2;\ 18 + 60\rho,\ \sigma_p^2), \tag{4.16}$$

where $\mathcal{N}(x; m, v)$ denotes the Gaussian density with mean $m$ and variance $v$ evaluated at $x$. The exact rows (4.8)–(4.9) are the limit $\sigma_p \to 0$, in which each density concentrates on its row. The derivation below keeps $\sigma_p$ finite, carries all three entries, and takes the limit at the end, where it is checked numerically; no row is used to eliminate an entry.

### 4.2.2 The prior density and Bayes' theorem

By equation (4.12) the prior density of the resistance is

$$p(\rho) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(\rho - \mu_0)^2}{2\sigma_0^2}\right] \propto \exp\left[-\frac{(\rho - 0.160)^2}{2(0.025)^2}\right]. \tag{4.17}$$

The prior of the state is the product of this density with the two physical densities (4.16), $p(\mathbf{z}) = p(T_1 \mid \rho)\, p(T_2 \mid \rho)\, p(\rho)$, the factorization of Chapter 3 read along the graph: each hall temperature depends on the resistance, and on nothing else. Bayes' theorem, derived in that chapter, gives the posterior of the state as

$$p(\mathbf{z} \mid y_T) = \frac{p(y_T \mid \mathbf{z})\, p(\mathbf{z})}{p(y_T)}, \qquad p(y_T) = \int p(y_T \mid \mathbf{u})\, p(\mathbf{u})\, \mathrm{d}\mathbf{u},$$

with $\mathbf{u}$ a variable of integration over the state. The denominator does not depend on $\mathbf{z}$, so

$$p(\mathbf{z} \mid y_T) \propto p(y_T \mid \mathbf{z})\, p(\mathbf{z}).$$

**The evidence of the model.** The denominator has a name of its own. Writing the model $M$ explicitly, it is the evidence, or marginal likelihood, of the model,

$$p(y_T \mid M) = \int p(y_T \mid \mathbf{z}, M)\, p(\mathbf{z} \mid M)\, \mathrm{d}\mathbf{z}\,: \tag{4.18}$$

the density of the observed reading averaged over the prior of the state. For the linear-Gaussian model it is the prior predictive density of the reading evaluated at the observed value (Section 4.4.2). It also follows from the objective and its curvature, which is how it is computed when no closed form exists. The integrand is $\exp[-\mathcal{L}(\mathbf{z})]$, with the objective $\mathcal{L}$ of equation (4.24), times the normalizing constants $(2\pi\sigma_k^2)^{-1/2}$ of the four rows; when the objective is quadratic about its minimum $\hat{\mathbf{z}}$ with matrix $\mathbf{J}^\mathsf{T}\mathbf{J}$, whose inverse is the covariance $\boldsymbol{\Sigma}$ of Section 4.2.5, the Gaussian integral over the three entries gives

$$p(y_T \mid M) = \frac{(2\pi)^{3/2}\, \det(\boldsymbol{\Sigma})^{1/2}}{(2\pi)^2\, \sigma_p^2\, \sigma_T\, \sigma_0}\, \mathrm{e}^{-\mathcal{L}(\hat{\mathbf{z}})}. \tag{4.19}$$

The two directions of the state that the physical rows fix have posterior widths proportional to $\sigma_p$, and they cancel the two physical rows' normalizing constants: $\det(\boldsymbol{\Sigma})^{1/2}/\sigma_p^2$ tends to the width of the posterior along the free direction. The result is the volume of the posterior, $\det(\boldsymbol{\Sigma})^{1/2}$, times the height of the integrand at the MAP state. The formula is exact when the objective is quadratic, as it is here; in general it is the Laplace approximation to the evidence.

The evidence normalizes the posterior, and its second use is to compare models. Alternatives for the hall include $M_1$, a single thermal resistance; $M_2$, two distinct heat-removal paths; and $M_3$, a model with a biased hall sensor. Each gives its own evidence for the same reading, and their ratio is a Bayes factor,

$$B_{21} = \frac{p(y_T \mid M_2)}{p(y_T \mid M_1)}, \tag{4.20}$$

which multiplies the prior odds of the two models to give their posterior odds, $\Pr(M_2 \mid y_T)/\Pr(M_1 \mid y_T) = B_{21} \cdot \Pr(M_2)/\Pr(M_1)$. That comparison is the category *hypothesis comparison*, and Chapter 27 develops it, the biased sensor among its drawings.

**The posterior up to its constant.** The data $\mathbf{y}$, the evidence in the sense of what is known about the hall, consists of the reading and the handbook prior; the likelihood involves only the reading, and the prior enters through $p(\mathbf{z})$, so $p(\mathbf{z} \mid \mathbf{y})$ and $p(\mathbf{z} \mid y_T)$ name the same density. Substituting equations (4.14), (4.16) and (4.17) and collecting the four exponents,

$$p(\mathbf{z} \mid \mathbf{y}) \propto \exp\left[-\frac{1}{2}\left(\frac{(T_1 - 18 - 30\rho)^2}{\sigma_p^2} + \frac{(T_2 - 18 - 60\rho)^2}{\sigma_p^2} + \frac{(T_1 - y_T)^2}{\sigma_T^2} + \frac{(\rho - \mu_0)^2}{\sigma_0^2}\right)\right]. \tag{4.21}$$

This is the posterior of the state up to its normalizing constant. It combines the physical relations, the reading, the reading's uncertainty and the prior knowledge in one probability model, one term per row. Every later result of the chapter is a property of this expression.

### 4.2.3 The residual vector, the Jacobian and the information matrix

The maximum a posteriori estimate is the state at which the posterior density is greatest,

$$\hat{\mathbf{z}} = \arg\max_{\mathbf{z}} \, p(\mathbf{z} \mid \mathbf{y}). \tag{4.22}$$

The logarithm is increasing, so maximizing the posterior is equivalent to minimizing its negative logarithm. Dropping the terms that do not depend on $\mathbf{z}$ (the normalizing constants of the four densities and $\ln p(y_T)$) leaves half the sum of the squared exponents of (4.21). Each exponent is a row divided by its standard deviation, and the four rows stack into one residual vector,

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T_1 - 18 - 30\rho)/\sigma_p \\ (T_2 - 18 - 60\rho)/\sigma_p \\ (T_1 - y_T)/\sigma_T \\ (\rho - \mu_0)/\sigma_0 \end{bmatrix} \begin{matrix} \text{physics, test hall} \\ \text{physics, full-load hall} \\ \text{reading} \\ \text{prior} \end{matrix} \tag{4.23}$$

whose entries are dimensionless, so that the objective is

$$\mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \qquad p(\mathbf{z} \mid \mathbf{y}) \propto \exp[-\mathcal{L}(\mathbf{z})]. \tag{4.24}$$

In the notation of the book it is $\mathcal{L} = \frac{1}{2}\sum_k w_k r_k^2$ over the unnormalized rows $r_1, r_2$ of equations (4.8)–(4.9), $T_1 - y_T$ and $\rho - \mu_0$, with weights $w_k = 1/\sigma_k^2$; the normalized rows carry their weights inside them. The first two terms penalize a state that unbalances a hall, in units of the model tolerance; the third, a state that disagrees with the reading, in units of the sensor spread; the fourth, a resistance far from its handbook value, in units of the prior spread. The weight of a physical row has a statistical meaning: $1/\sigma_p^2$ is the inverse variance of a model discrepancy, the amount by which the steady-state balance may fail.

**The Jacobian.** The derivatives of the four rows with respect to the three entries $(\rho, T_1, T_2)$ form the $4 \times 3$ Jacobian $\mathbf{J} = \partial\mathbf{r}/\partial\mathbf{z}$, one line per row and one column per entry:

$$\mathbf{J} = \begin{bmatrix} -30/\sigma_p & 1/\sigma_p & 0 \\ -60/\sigma_p & 0 & 1/\sigma_p \\ 0 & 1/\sigma_T & 0 \\ 1/\sigma_0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -3000 & 100 & 0 \\ -6000 & 0 & 100 \\ 0 & 2 & 0 \\ 40 & 0 & 0 \end{bmatrix} \quad (\sigma_p = 0.01\,\mathrm{K}), \tag{4.25}$$

with the first column in kW/K and the other two in $\mathrm{K}^{-1}$. The entry $J_{ki}$ is the change in row $k$, counted in units of that row's standard deviation, per unit change of entry $i$. Its nonzero entries sit where the pattern (4.13) of the graph places them. The entries $2 = 1/0.5$ and $40 = 1/0.025$ are the normalizing factors of the reading and the prior. Every row is linear in the state, so $\mathbf{J}$ is the same at every linearization point. The tolerance $\sigma_p = 0.01\,\mathrm{K}$, one-fiftieth of the sensor's spread, is tight enough that the answers no longer move (the limit is checked in Section 4.4.5) and loose enough, within two decades of $\sigma_T$, to keep the normal equations conditioned well enough for double precision.

**The information matrix.** For rows normalized by their standard deviations the information matrix is $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$, equation (3.26). Written with the unnormalized rows, whose Jacobian is $\tilde{\mathbf{J}}$, and the diagonal weight matrix $\mathbf{W} = \mathrm{diag}(1/\sigma_p^2,\ 1/\sigma_p^2,\ 1/\sigma_T^2,\ 1/\sigma_0^2)$, it is the same matrix in the form used throughout the book,

$$\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J} = \tilde{\mathbf{J}}^\mathsf{T}\mathbf{W}\tilde{\mathbf{J}} = \mathbf{J}_\mathrm{d}^\mathsf{T}\mathbf{W}_\mathrm{d}\mathbf{J}_\mathrm{d} + \mathbf{\Lambda}_\mathrm{prior}, \tag{4.26}$$

where the last form keeps the prior rows apart: $\mathbf{J}_\mathrm{d}$ and $\mathbf{W}_\mathrm{d}$ hold the physical and reading rows, and $\mathbf{\Lambda}_\mathrm{prior}$, here $\mathrm{diag}(1/\sigma_0^2, 0, 0)$, is the prior precision. For rows that are nonlinear in the state, equation (4.26) is the Gauss–Newton approximation to the Hessian of $\mathcal{L}$: it drops the curvature term $\sum_k r_k\, \partial^2 r_k/\partial\mathbf{z}\,\partial\mathbf{z}^\mathsf{T}$, the second derivatives of the rows weighted by their residuals, which vanishes for linear rows. Writing $\mathbf{j}_k$ for the $k$-th line of $\mathbf{J}$, the product is a sum of one outer product per row, so the information of independent rows adds, as equation (3.27) states for priors and readings; with the numbers of equation (4.25),

$$\mathbf{J}^\mathsf{T}\mathbf{J} = \sum_{k=1}^{4} \mathbf{j}_k^\mathsf{T}\mathbf{j}_k = \begin{bmatrix} 45\,001\,600 & -300\,000 & -600\,000 \\ -300\,000 & 10\,004 & 0 \\ -600\,000 & 0 & 10\,000 \end{bmatrix}. \tag{4.27}$$

The entry in the corner is $3000^2 + 6000^2 + 40^2$: the two physical rows and the prior. The reading contributes only the $4 = 1/\sigma_T^2$ of the $(T_1, T_1)$ entry. The reading and prior rows alone would give $\mathrm{diag}(1600, 4, 0)$: no information at all would reach $T_2$, and none would pass from the reading to $\rho$. It is the off-diagonal entries of the physical rows that carry the reading at $T_1$ to $\rho$ and from $\rho$ to $T_2$. The matrix does not contain the reading's value $y_T$, only its standard deviation: the information is fixed by the design of the measurement and by the prior, not by the value observed. Its eigenvalues are 1.155, $1.0003 \times 10^4$ and $4.5012 \times 10^7$, all positive, with a condition number of $3.9 \times 10^7$ that grows as $1/\sigma_p^2$.

### 4.2.4 The normal equations and the MAP state

**The stationary condition.** By the chain rule the gradient of the objective (4.24) is

$$\nabla\mathcal{L}(\mathbf{z}) = \mathbf{J}^\mathsf{T}\mathbf{r}(\mathbf{z}), \tag{4.28}$$

and the MAP state is the point where it vanishes, $\mathbf{J}^\mathsf{T}\mathbf{r}(\hat{\mathbf{z}}) = \mathbf{0}$: three equations, one per entry.

**The Gauss–Newton step.** The step that sets the gradient of the linearized rows to zero solves the normal equations of equation (3.25),

$$(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \qquad \mathbf{z}^{(k+1)} = \mathbf{z}^{(k)} + \Delta\mathbf{z}, \tag{4.29}$$

the Gauss–Newton iteration. The rows of this example are linear, so the linearization is exact and one step lands on $\hat{\mathbf{z}}$ from any starting state.

**The free direction.** The two physical lines of $\mathbf{J}$ annihilate one direction of the state,

$$\mathbf{J}_{\text{phys}}\,\mathbf{v} = \mathbf{0}, \qquad \mathbf{J}_{\text{phys}} = \begin{bmatrix} -3000 & 100 & 0 \\ -6000 & 0 & 100 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 1 \\ 30 \\ 60 \end{bmatrix}, \tag{4.30}$$

the free direction: a change of $\rho$ with $T_1$ and $T_2$ following by $Q_1 = 30\,\text{kW}$ and $Q_2 = 60\,\text{kW}$ times that change, which unbalances neither hall. A pull of the reading on $T_1$ is answered by a move along $\mathbf{v}$, because any move off it would cost the physical rows far more than it gains on the reading; from a state that balances both halls, the step of the normal equations therefore lies, to within the tolerance, along $\mathbf{v}$.

### 4.2.5 The posterior covariance from the information matrix

The point estimate does not answer the release question, which depends on how uncertain the state remains.

**The covariance.** The rows are linear and Gaussian, so by equation (3.26) the posterior is exactly $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ with

$$\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1},$$

the inverse of the matrix (4.27), with rows and columns ordered $(\rho, T_1, T_2)$, so that its entries carry $(\text{K/kW})^2$, $\text{K}^2$, and K·K/kW for the products. A larger information means a narrower posterior: the covariance is the inverse of the steepness of the objective.

**The marginals.** The estimation questions ask for entries. The marginals of a Gaussian are read off its mean and covariance,

$$\rho \mid \mathbf{y} \sim \mathcal{N}(\hat{z}_1,\ \Sigma_{11}), \qquad T_1 \mid \mathbf{y} \sim \mathcal{N}(\hat{z}_2,\ \Sigma_{22}), \qquad T_2 \mid \mathbf{y} \sim \mathcal{N}(\hat{z}_3,\ \Sigma_{33}), \tag{4.31}$$

so the resistance is $\hat{\rho} = \hat{z}_1$ with standard deviation $\sqrt{\Sigma_{11}}$, the test temperature $\hat{z}_2$ with $\sqrt{\Sigma_{22}}$, and the full-load temperature, the entry the threshold question asks about, $\hat{T}_2 = \hat{z}_3$ with $\sqrt{\Sigma_{33}}$. How the entries move together is read off the correlations $\Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$, which the physical rows drive toward one.

**Coordinates along and across the physical rows.** The shape of $\boldsymbol{\Sigma}$, and its limit as the tolerance shrinks, are read by measuring the state along the free direction and across the two physical rows. Write

$$\mathbf{z} = \mathbf{c} + \mathbf{B}\mathbf{u}, \qquad \mathbf{u} = \begin{bmatrix} \rho \\ d_1 \\ d_2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 30 & 1 & 0 \\ 60 & 0 & 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 0 \\ 18 \\ 18 \end{bmatrix}, \tag{4.32}$$

where $d_1 = T_1 - 18 - 30\rho$ and $d_2 = T_2 - 18 - 60\rho$ are the two hall imbalances in kelvin and the first column of $\mathbf{B}$ is $\mathbf{v}$. The rows are linear in $\mathbf{u}$ with Jacobian $\mathbf{JB}$, whose columns are $\mathbf{Jv}$, $\mathbf{Je}_2$ and $\mathbf{Je}_3$,

$$\mathbf{JB} = \begin{bmatrix} 0 & 1/\sigma_p & 0 \\ 0 & 0 & 1/\sigma_p \\ 30/\sigma_T & 1/\sigma_T & 0 \\ 1/\sigma_0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 100 & 0 \\ 0 & 0 & 100 \\ 60 & 2 & 0 \\ 40 & 0 & 0 \end{bmatrix}.$$

The first column, $\mathbf{Jv} = (0,\ 0,\ 60,\ 40)^{\mathsf{T}}$, is seen only by the reading (with slope $2 \times 30 = 60$) and the prior (slope 40); a move along $\mathbf{v}$ costs the information $\|\mathbf{Jv}\|^2 = 60^2 + 40^2 = 5200\,(\mathrm{kW/K})^2$. The other two columns carry the entries $1/\sigma_p$, so every direction that unbalances a hall costs about $1/\sigma_p^2 = 10^4\,\mathrm{K}^{-2}$: a tiny tolerance leaves the state free to move cheaply only along $\mathbf{v}$, a narrow ridge in the objective. The third column is orthogonal to the other two, so the information matrix in these coordinates is block diagonal,

$$(\mathbf{JB})^{\mathsf{T}}(\mathbf{JB}) = \begin{bmatrix} 1/\sigma_0^2 + 30^2/\sigma_T^2 & 30/\sigma_T^2 & 0 \\ 30/\sigma_T^2 & 1/\sigma_p^2 + 1/\sigma_T^2 & 0 \\ 0 & 0 & 1/\sigma_p^2 \end{bmatrix} = \begin{bmatrix} 5200 & 120 & 0 \\ 120 & 10\,004 & 0 \\ 0 & 0 & 10\,000 \end{bmatrix},$$

and its inverse, the covariance $\mathbf{\Sigma}_u$ of $\mathbf{u}$, is

$$\mathbf{\Sigma}_u = \begin{bmatrix} 1.9236 \times 10^{-4} & -2.3074 \times 10^{-6} & 0 \\ -2.3074 \times 10^{-6} & 9.9988 \times 10^{-5} & 0 \\ 0 & 0 & 1.0000 \times 10^{-4} \end{bmatrix} \quad (\sigma_p = 0.01\,\mathrm{K}).$$

Three results follow. First, the full-load imbalance $d_2$ is seen by its own row only: $(\mathbf{\Sigma}_u)_{33} = \sigma_p^2$, the variance that row allows, and it is uncorrelated with $\rho$ and $d_1$. Second, by the block-inversion formula of Section 3.5.3, the first diagonal entry of $\mathbf{\Sigma}_u$ is the inverse of the Schur complement of the $(d_1, d_1)$ entry in the upper $2 \times 2$ block,

$$\Sigma_{11} = (\mathbf{\Sigma}_u)_{11} = \left[ \frac{1}{\sigma_0^2} + \frac{30^2}{\sigma_T^2} - \frac{(30/\sigma_T^2)^2}{1/\sigma_p^2 + 1/\sigma_T^2} \right]^{-1} = \left[ \frac{1}{\sigma_0^2} + \frac{30^2}{\sigma_T^2 + \sigma_p^2} \right]^{-1}, \tag{4.33}$$

the first row of $\mathbf{B}$ being $\mathbf{e}_1^{\mathsf{T}}$; a tolerance acts on the reading as extra sensor variance and widens the spread of $\rho$. Third, the covariance of the state is $\mathbf{\Sigma} = \mathbf{B}\mathbf{\Sigma}_u\mathbf{B}^{\mathsf{T}}$; its third diagonal entry, with $T_2 = 18 + 60\rho + d_2$ and $d_2$ uncorrelated with $\rho$, is

$$\mathrm{Var}(T_2 \mid \mathbf{y}) = \Sigma_{33} = \mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g} + \sigma_p^2 = 60^2\,\Sigma_{11} + \sigma_p^2, \qquad \mathbf{g} = (60,\ 0,\ 0)^{\mathsf{T}}, \tag{4.34}$$

with $\mathbf{g}$ the gradient of the prediction $18 + 60\rho$. A model discrepancy therefore enters the answer twice: it widens $\Sigma_{11}$, and it adds its own variance at the node that has no reading.

**The limit of exact rows.** As $\sigma_p \to 0$, $(\mathbf{\Sigma}_u)_{22}$ and $(\mathbf{\Sigma}_u)_{33}$, both at most $\sigma_p^2$, vanish and $(\mathbf{\Sigma}_u)_{11}$ tends to $1/5200$ by equation (4.33), so $\mathbf{\Sigma}_u \to \mathrm{diag}(1/5200,\ 0,\ 0)$ and

$$\mathbf{\Sigma} = \mathbf{B}\mathbf{\Sigma}_u\mathbf{B}^{\mathsf{T}} \ \longrightarrow\ \frac{\mathbf{B}\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}\mathbf{B}^{\mathsf{T}}}{5200} = \frac{\mathbf{v}\mathbf{v}^{\mathsf{T}}}{5200} = \frac{1}{5200}\begin{bmatrix} 1 & 30 & 60 \\ 30 & 900 & 1800 \\ 60 & 1800 & 3600 \end{bmatrix}, \tag{4.35}$$

of rank one. The matrix has two zero eigenvalues and no ordinary inverse, and the posterior correlation of $T_1$ and $T_2$ is exactly one. This is not a failure of the inference; it is the consequence of exact physical rows, which leave the state one degree of freedom. The state-vector form needs no special treatment of it: the normal equations are solved at a finite tolerance, where $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ is invertible, and the limit is derived in the coordinates of equation (4.32) and read from the sequence of tolerances. The distinction between exact constraints and uncertain physical relations matters whenever a Hessian and a covariance are constructed over the full state, and the tolerance is the device that keeps both defined.

**The objective along the free direction.** On the line $\mathbf{z} = \hat{\mathbf{z}} + t\,\mathbf{v}$ through the MAP state, in the limit of exact rows, the physical rows keep their values at $\hat{\mathbf{z}}$ (zero in the limit) and only the reading and prior rows change, as $r_3(\hat{\mathbf{z}}) + 60t$ and $r_4(\hat{\mathbf{z}}) + 40t$, the slopes being the entries of $\mathbf{J}\mathbf{v} = (0,\ 0,\ 60,\ 40)^{\mathsf{T}}$. The objective on the line is the reading term $\frac{1}{2}(r_3(\hat{\mathbf{z}}) + 60t)^2$ plus the prior term $\frac{1}{2}(r_4(\hat{\mathbf{z}}) + 40t)^2$. At the minimum the slopes of the two terms cancel, $\mathbf{v}^{\mathsf{T}}\mathbf{J}^{\mathsf{T}}\mathbf{r}(\hat{\mathbf{z}}) = 60\,r_3 + 40\,r_4 = 0$: the reading pulls with weight 60 and the prior with weight 40. The gradient vanishes at $\hat{\mathbf{z}}$, so the objective on the line is

$$\mathcal{L}(\hat{\mathbf{z}}+t\mathbf{v}) = \mathcal{L}(\hat{\mathbf{z}})+\tfrac{1}{2}\,t^2\,\|\mathbf{J}\mathbf{v}\|^2, \qquad \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{L}(\hat{\mathbf{z}}+t\mathbf{v}) = \mathbf{v}^{\mathsf{T}}\mathbf{J}^{\mathsf{T}}\mathbf{J}\,\mathbf{v} = 60^2+40^2 = 5200\,(\mathrm{kW/K})^2, \tag{4.36}$$

a parabola whose curvature is the information along $\mathbf{v}$, equal to $1/\Sigma_{11}$ in the limit; the rise of $\frac{1}{2}$ is reached at $t = \pm 1/\sqrt{5200} = \pm\sqrt{\Sigma_{11}}$.

*Remark* 4.2. Substituting the exact rows $T_1 = 18 + 30\rho$ and $T_2 = 18 + 60\rho$ into the objective gives the same curve as a function of the resistance alone,

$$\mathcal{L}(\rho) = \frac{1}{2}\left[\frac{(\rho-\mu_0)^2}{\sigma_0^2} + \frac{(y_T - 18 - 30\rho)^2}{\sigma_T^2}\right], \tag{4.37}$$

whose second derivative is the 5200 of equation (4.36) and whose minimum is $\hat{\rho} = 880/5200$.

### 4.2.6 The Gaussian form of the posterior

In this example the prior is Gaussian, the reading error and the model tolerance are Gaussian, and every row is linear in the state. The negative log posterior is therefore exactly quadratic in $\mathbf{z}$, the curvature term $\sum_k r_k\,\partial^2 r_k/\partial\mathbf{z}\,\partial\mathbf{z}^{\mathsf{T}}$ is zero, the Gauss–Newton matrix $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ is the exact Hessian, and its inverse is the exact posterior covariance under the stated model. Nothing in equations (4.29)–(4.31) is an approximation; the only limit taken, $\sigma_p \to 0$, is derived in equation (4.35).

For a nonlinear physical model, for instance one written in the log-conductance $g$, the negative log posterior is not quadratic. The same construction is then made locally: the objective is expanded to second order about the MAP state,

$$\mathcal{L}(\mathbf{z}) \approx \mathcal{L}(\hat{\mathbf{z}}) + \tfrac{1}{2}(\mathbf{z}-\hat{\mathbf{z}})^{\mathsf{T}}\mathbf{H}(\mathbf{z}-\hat{\mathbf{z}}), \qquad \mathbf{H} = \left.\frac{\partial^2\mathcal{L}}{\partial\mathbf{z}\,\partial\mathbf{z}^{\mathsf{T}}}\right|_{\hat{\mathbf{z}}} = \mathbf{J}^{\mathsf{T}}\mathbf{J} + \sum_k r_k \left.\frac{\partial^2 r_k}{\partial\mathbf{z}\,\partial\mathbf{z}^{\mathsf{T}}}\right|_{\hat{\mathbf{z}}},$$

with no first-order term because the gradient vanishes at the minimum, and the posterior is replaced by the Gaussian that this expansion defines, the Laplace approximation:

$$\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}},\ \mathbf{H}^{-1}). \tag{4.38}$$

It holds when the posterior is close to Gaussian over its own width and when $\mathbf{H}$ is positive definite in the chosen free coordinates. The distinction matters throughout the book: the calculation of this chapter is exact, while a Laplace approximation in general is not, and its adequacy is one of the checks of step 5.

### 4.2.7 Rows that are not Gaussian

Every row of the hall is Gaussian, and its term in the objective is half its squared residual. Chapter 2's trip log already has a row of another kind, a count of loop trips in a number of windows. The construction above carries over to it unchanged once the objective is defined by the posterior rather than by a sum of squares.

**The objective.** The objective is the negative logarithm of the posterior up to a constant. Each row, a likelihood factor or a prior factor, contributes one term, its negative log-probability:

$$\mathcal{L}(\mathbf{z}) = \sum_k \mathcal{L}_k(\mathbf{z}), \qquad \mathcal{L}_k(\mathbf{z}) = -\ln p_k(\cdot \mid \mathbf{z}), \qquad p(\mathbf{z} \mid \mathbf{y}) \propto \exp[-\mathcal{L}(\mathbf{z})]. \tag{4.39}$$

A Gaussian row is the case $\mathcal{L}_k = \frac{1}{2} r_k^2$, and the objective (4.24) of the hall is this sum for its four rows.

**What the fit uses.** The MAP state, the Newton step and the Laplace covariance use the objective only through its gradient and its Hessian, and both are sums over the rows:

$$\nabla\mathcal{L} = \sum_k \nabla\mathcal{L}_k, \qquad \mathbf{H} = \sum_k \nabla^2\mathcal{L}_k, \qquad \mathbf{H}\,\Delta\mathbf{z} = -\nabla\mathcal{L}, \qquad \mathbf{\Sigma} = \mathbf{H}(\hat{\mathbf{z}})^{-1}. \tag{4.40}$$

The value of $\mathcal{L}$ enters only the evidence, equation (4.19). When the state has one entry $\xi$, the gradient is the slope $\mathcal{L}'(\xi)$, the Hessian the curvature $\mathcal{L}''(\xi)$, the step $\Delta\xi = -\mathcal{L}'/\mathcal{L}''$ and the variance $1/\mathcal{L}''(\hat{\xi})$; nothing else changes.

**The Gaussian row.** For $\mathcal{L}_k = \frac{1}{2} r_k^2$ with Jacobian line $\mathbf{j}_k = \partial r_k / \partial \mathbf{z}^\mathsf{T}$, the chain rule gives

$$\nabla\mathcal{L}_k = \mathbf{j}_k^\mathsf{T} r_k, \qquad \nabla^2\mathcal{L}_k = \mathbf{j}_k^\mathsf{T}\mathbf{j}_k + r_k \frac{\partial^2 r_k}{\partial \mathbf{z}\,\partial \mathbf{z}^\mathsf{T}}. \tag{4.41}$$

Summed over the rows these are $\mathbf{J}^\mathsf{T}\mathbf{r}$ and $\mathbf{J}^\mathsf{T}\mathbf{J}$ plus the curvature term of Section 4.2.6.

**The count row.** A record of $n_k$ events (the count $k$ of Chapter 2) whose expected number is $\mu_k(\mathbf{z})$, Poisson in windows, has the term $\mathcal{L}_k = \mu_k - n_k \ln \mu_k$. Its gradient and Hessian are

$$\nabla\mathcal{L}_k = (\mu_k - n_k)\,\nabla \ln \mu_k, \qquad \nabla^2\mathcal{L}_k = \mu_k\,\nabla \ln \mu_k\,\nabla \ln \mu_k^\mathsf{T} + (\mu_k - n_k)\,\nabla^2 \ln \mu_k. \tag{4.42}$$

The first term of the Hessian is the product of one line with itself, so it is $\mathbf{j}_k^\mathsf{T}\mathbf{j}_k$ for the line and the residual

$$\mathbf{j}_k = \sqrt{\mu_k}\,\nabla \ln \mu_k^\mathsf{T}, \qquad r_k = \frac{\mu_k - n_k}{\sqrt{\mu_k}}, \tag{4.43}$$

which also give $\mathbf{j}_k^\mathsf{T} r_k = \nabla\mathcal{L}_k$. This pair, the *working residual* and the *working Jacobian line*, is how a count row joins the Jacobian beside Gaussian rows: one more line of $\mathbf{J}$, one more entry of $\mathbf{r}$, and the same products (4.40). The remaining term $(\mu_k - n_k)\,\nabla^2 \ln \mu_k$ plays the part of $r_k\,\partial^2 r_k$ in a Gaussian row. It vanishes when $\ln \mu_k$ is linear in the state, which is the case when the state carries the log-rate, the canonical parameter of the Poisson distribution.

The working pair reproduces the row's gradient and the Gauss–Newton part of its Hessian, and nothing more. Its half square is not the row's term, $\frac{1}{2} r_k^2 \neq \mathcal{L}_k$, and $\mathbf{j}_k$ is not the derivative of $r_k$ except where $r_k = 0$; the fit never uses either, because it uses only the products (4.40). For a Gaussian row that is linear in the state, the working pair is the row's own residual and Jacobian line, which is why the two forms are the same in the hall.

**The information and its curvature term.** With every row in this form, the Hessian is

$$\mathbf{H} = \mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{C}, \qquad \mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}, \tag{4.44}$$

with $\mathbf{C}$ the sum of the rows' remaining terms. The matrix $\mathbf{\Lambda}$ is the information; it is the average of $\mathbf{H}$ over the records the model could have produced, the *expected information*. The two are equal when $\mathbf{C} = \mathbf{0}$: for linear Gaussian rows, and for count rows carried on their canonical parameter. Where $\mathbf{C}$ is not zero, the chapters name it, keep it in the Newton step and report the covariance from $\mathbf{H}$. Once $\hat{\mathbf{z}}$ and $\mathbf{\Sigma}$ are in hand, the kind of row no longer matters: every answer of step 4 is read from them in the same way.

### 4.2.8 The reduced state

The full state carries every free quantity of the graph, and in the limit of exact rows the physical rows make some of its entries functions of the others; in the hall, two of the three entries follow from the third. The same posterior can then be written over the entries that remain free, the reduced state, as a projection of the rows onto the null space of the physical rows; the two-unit hall of Chapter 11 and the larger plants after it are fitted in this form, so the derivation below is written for any plant, with the hall as its instance.

**The null space of the physical rows.** Let the full state $\mathbf{z}$ have $n$ entries and the graph $m$ physical rows, with Jacobian $\mathbf{A} = \mathbf{J}_{\text{phys}}$, an $m \times n$ matrix. Split the full state into $n_{\text{f}} = n - m$ free entries $\mathbf{z}_{\text{f}}$ and $m$ dependent entries $\mathbf{z}_{\text{d}}$, and $\mathbf{A}$ into the matching columns, $\mathbf{A} = [\,\mathbf{A}_{\text{free}}\ \ \mathbf{A}_{\text{dep}}\,]$, an $m \times n_{\text{f}}$ and an $m \times m$ block (the entries ordered free first). The split is a choice: any split serves for which $\mathbf{A}_{\text{dep}}$ is square and invertible, that is, for which the physical rows can be solved for the dependent entries at given free ones. A move $\Delta\mathbf{z}$ of the full state keeps the linearized balances when $\mathbf{A}\,\Delta\mathbf{z} = \mathbf{0}$, that is $\mathbf{A}_{\text{free}}\,\Delta\mathbf{z}_{\text{f}} + \mathbf{A}_{\text{dep}}\,\Delta\mathbf{z}_{\text{d}} = \mathbf{0}$, so

$$\Delta\mathbf{z}_{\text{d}} = \mathbf{N}_{\text{b}}\,\Delta\mathbf{z}_{\text{f}}, \qquad \mathbf{N}_{\text{b}} = -\mathbf{A}_{\text{dep}}^{-1}\mathbf{A}_{\text{free}}, \qquad \mathbf{N} = \begin{bmatrix} \mathbf{I}_{n_{\text{f}}} \\ \mathbf{N}_{\text{b}} \end{bmatrix}, \qquad \mathbf{A}\mathbf{N} = \mathbf{0}. \tag{4.45}$$

Each physical row's tolerance divides both of its blocks and cancels in $\mathbf{N}_{\text{b}}$. The $n_{\text{f}}$ columns of the $n \times n_{\text{f}}$ matrix $\mathbf{N}$ are independent, through its identity block, and $\mathbf{A}$ has rank $m$, so they span the null space of the physical rows, the tangent of the set on which the balances hold; each column moves one free entry by one unit and the dependent entries as the linearized balances require. In the hall $n = 3$ and $m = 2$; with $\mathbf{z}_{\text{f}} = \rho$ and $\mathbf{z}_{\text{d}} = (T_1, T_2)^{\mathsf{T}}$,

$$\mathbf{A}_{\text{free}} = \begin{bmatrix} -Q_1/\sigma_p \\ -Q_2/\sigma_p \end{bmatrix}, \qquad \mathbf{A}_{\text{dep}} = \begin{bmatrix} 1/\sigma_p & 0 \\ 0 & 1/\sigma_p \end{bmatrix}, \qquad \mathbf{N}_{\text{b}} = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}, \qquad \mathbf{N} = \begin{bmatrix} 1 \\ 30 \\ 60 \end{bmatrix},$$

the entries of $\mathbf{N}_{\text{b}}$ the loads in kW: a unit change of resistance raises each hall temperature by its load. The single column of $\mathbf{N}$ is the free direction $\mathbf{v}$ of equation (4.30). Taking $T_1$ as the free entry would serve as well, since the block of $(\rho, T_2)$ is invertible too; the resistance is chosen because the prior acts on it and each hall temperature follows from it by its row of $\mathbf{N}_{\text{b}}$.

**The rows of the reduced state.** With exact balances the full state lies on the set where the physical rows vanish, and the reduced state $\mathbf{z}_{\text{f}}$ names a point of that set; the point itself is $\mathbf{z}_{\text{full}}(\mathbf{z}_{\text{f}}) = (\mathbf{z}_{\text{f}},\ \mathbf{z}_{\text{d}}(\mathbf{z}_{\text{f}}))$, with $\mathbf{z}_{\text{d}}(\mathbf{z}_{\text{f}})$ the solution of the physical rows at the given free entries, and its derivative with respect to $\mathbf{z}_{\text{f}}$ is $\mathbf{N}$. Write $\mathbf{r}_{\text{rows}}$ for the evidence rows, the readings and the priors, as functions of the full state. The rows of the reduced state are the evidence rows at the point of the set,

$$\mathbf{r}(\mathbf{z}_{\text{f}}) = \mathbf{r}_{\text{rows}}\big(\mathbf{z}_{\text{full}}(\mathbf{z}_{\text{f}})\big), \tag{4.46}$$

and the objective on the reduced state is the full objective (4.24) on the set, where its physical rows contribute nothing. In the hall, with $\mathbf{z}_{\text{full}}(\rho) = (\rho,\ T_c + Q_1\rho,\ T_c + Q_2\rho)^{\mathsf{T}}$, the rows are

$$\mathbf{r}(\rho) = \begin{bmatrix} (T_c + Q_1\rho - y_T)/\sigma_T \\ (\rho - \mu_0)/\sigma_0 \end{bmatrix} \begin{matrix} \text{reading} \\ \text{prior} \end{matrix},$$

and the objective is the curve (4.37), reached here as a projection and a limit rather than by a substitution made by hand.

**The reduced Jacobian.** A move $\Delta\mathbf{z}_\mathrm{f}$ of the reduced state moves the point of the set by $\mathbf{N}\,\Delta\mathbf{z}_\mathrm{f}$, so the chain rule gives the Jacobian of the reduced rows as the Jacobian of the evidence rows projected onto the null space,

$$\mathbf{J}_\mathrm{red} = \mathbf{J}_\mathrm{rows}\mathbf{N}, \tag{4.47}$$

where $\mathbf{J}_\mathrm{rows} = \partial\mathbf{r}_\mathrm{rows}/\partial\mathbf{z}$ is the Jacobian of the evidence rows (readings and priors) with respect to the full state, the lines of $\mathbf{J}$ that are not physical rows. A reading whose row touches only dependent entries sees the free entries through $\mathbf{N}_\mathrm{b}$. In the hall $\mathbf{J}_\mathrm{rows}$ is the reading and prior lines of equation (4.25),

$$\mathbf{J}_\mathrm{rows} = \begin{bmatrix} 0 & 1/\sigma_T & 0 \\ 1/\sigma_0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{J}_\mathrm{red} = \begin{bmatrix} Q_1/\sigma_T \\ 1/\sigma_0 \end{bmatrix}.$$

The reading, whose row in the full state touches only $T_1$, sees the resistance on the reduced state with the slope $Q_1/\sigma_T$: its entry is nonzero, so the reading alone determines the one free entry (the hall is observable from its one sensor). The information of the reduced state is $\mathbf{J}_\mathrm{red}^\mathsf{T}\mathbf{J}_\mathrm{red} = Q_1^2/\sigma_T^2 + 1/\sigma_0^2$, the reading's term and the prior's, equal to $\|\mathbf{J}\mathbf{v}\|^2$ of equation (4.36).

**Nonlinear physical rows.** When the physical rows are nonlinear in the state, as they are when a conductance is carried as a log-derate or a recirculation fraction multiplies a temperature difference, the same objects are evaluated point by point. $\mathbf{z}_\mathrm{full}(\mathbf{z}_\mathrm{f})$ is the solution of the physical rows at the given free entries, found by Newton's method on those rows alone; $\mathbf{N}(\mathbf{z}_\mathrm{f})$, from the blocks of $\mathbf{A}$ at that solution, and $\mathbf{J}_\mathrm{red}(\mathbf{z}_\mathrm{f}) = \mathbf{J}_\mathrm{rows}\,\mathbf{N}(\mathbf{z}_\mathrm{f})$ are evaluated there and change with the free entries. The Gauss–Newton iteration (4.29) runs on the reduced rows, $(\mathbf{J}_\mathrm{red}^\mathsf{T}\mathbf{J}_\mathrm{red})\,\Delta\mathbf{z}_\mathrm{f} = -\mathbf{J}_\mathrm{red}^\mathsf{T}\mathbf{r}(\mathbf{z}_\mathrm{f})$, each iterate solving the physical rows again. At the MAP state, $\mathbf{\Sigma}_\mathrm{red} = (\mathbf{J}_\mathrm{red}^\mathsf{T}\mathbf{J}_\mathrm{red})^{-1}$ is the Gauss–Newton form of the Laplace covariance (4.38); the term it drops holds the second derivatives of the reduced rows, which now include the curvature of the set, and each chapter checks it numerically. The hall's rows are linear, $\mathbf{N}$ is constant, one step is exact and $\mathbf{\Sigma}_\mathrm{red}$ is the exact covariance.

**The reduced covariance as the limit.** The covariance of the reduced state is $\mathbf{\Sigma}_\mathrm{red} = (\mathbf{J}_\mathrm{red}^\mathsf{T}\mathbf{J}_\mathrm{red})^{-1}$, and it is the free-entry block $\mathbf{\Sigma}_\mathrm{ff}$ of the full $\mathbf{\Sigma}$ in the limit of exact rows. Write a move of the full state in the coordinates $\Delta\mathbf{z} = \mathbf{B}\,\Delta\mathbf{u}$ with $\mathbf{B} = [\,\mathbf{N}\;\;\mathbf{B}_\mathrm{d}\,]$, where $\mathbf{B}_\mathrm{d} = [\,\mathbf{0}\;\;\mathbf{I}_m\,]^\mathsf{T}$ is the $n \times m$ matrix that moves only the dependent entries. The square matrix $\mathbf{B}$ is invertible and its first block row is $[\,\mathbf{I}_{n_\mathrm{f}}\;\;\mathbf{0}\,]$, so the first $n_\mathrm{f}$ coordinates are the move of the free entries; in the hall $\mathbf{B}$ is the matrix of equation (4.32). Since $\mathbf{A}\mathbf{N} = \mathbf{0}$ and $\mathbf{A}\mathbf{B}_\mathrm{d} = \mathbf{A}_\mathrm{dep}$, the Jacobian in these coordinates, physical rows first, and its information matrix are

$$\mathbf{J}\,[\,\mathbf{N}\;\;\mathbf{B}_\mathrm{d}\,] = \begin{bmatrix} \mathbf{0} & \mathbf{A}_\mathrm{dep} \\ \mathbf{J}_\mathrm{red} & \mathbf{J}_\mathrm{rows}\mathbf{B}_\mathrm{d} \end{bmatrix}, \qquad \begin{bmatrix} \mathbf{J}_\mathrm{red}^\mathsf{T}\mathbf{J}_\mathrm{red} & \mathbf{E} \\ \mathbf{E}^\mathsf{T} & \mathbf{F} + \mathbf{A}_\mathrm{dep}^\mathsf{T}\mathbf{A}_\mathrm{dep} \end{bmatrix},$$

with $\mathbf{E} = \mathbf{J}_\mathrm{red}^\mathsf{T}\mathbf{J}_\mathrm{rows}\mathbf{B}_\mathrm{d}$ and $\mathbf{F} = \mathbf{B}_\mathrm{d}^\mathsf{T}\mathbf{J}_\mathrm{rows}^\mathsf{T}\mathbf{J}_\mathrm{rows}\mathbf{B}_\mathrm{d}$. By the block-inversion formula of Section 3.5.3 the free-entry block of the covariance in these coordinates, which is $\mathbf{\Sigma}_\mathrm{ff}$ because the first block row of $\mathbf{B}$ is $[\,\mathbf{I}_{n_\mathrm{f}}\;\;\mathbf{0}\,]$, is the inverse of the Schur complement of the lower block, and that block grows without bound as the tolerances shrink: each line of $\mathbf{A}_\mathrm{dep}$ carries its row's $1/\sigma_p$, so with a common tolerance $\mathbf{A}_\mathrm{dep}^\mathsf{T}\mathbf{A}_\mathrm{dep} = \sigma_p^{-2}\,\tilde{\mathbf{A}}_\mathrm{dep}^\mathsf{T}\tilde{\mathbf{A}}_\mathrm{dep}$, with $\tilde{\mathbf{A}}_\mathrm{dep}$ the invertible block of the unnormalized rows:

$$\mathbf{\Sigma}_\mathrm{ff} = \Big[\mathbf{J}_\mathrm{red}^\mathsf{T}\mathbf{J}_\mathrm{red} - \mathbf{E}\big(\mathbf{F} + \mathbf{A}_\mathrm{dep}^\mathsf{T}\mathbf{A}_\mathrm{dep}\big)^{-1}\mathbf{E}^\mathsf{T}\Big]^{-1} \;\longrightarrow\; \big(\mathbf{J}_\mathrm{red}^\mathsf{T}\mathbf{J}_\mathrm{red}\big)^{-1} = \mathbf{\Sigma}_\mathrm{red} \qquad (\sigma_p \to 0). \tag{4.48}$$

The correction the bracket subtracts is of order $\sigma_p^2$; for nonlinear rows the equality holds with every block evaluated at the point of linearization, the MAP state. In the hall $\mathbf{\Sigma}_\mathrm{ff}$ is the scalar

$\Sigma_{11}$, $\mathbf{E} = (Q_1/\sigma_T^2,\ 0)$, $\mathbf{F} = \mathrm{diag}(1/\sigma_T^2,\ 0)$ and $\mathbf{A}_{\mathrm{dep}}^{\mathsf{T}}\mathbf{A}_{\mathrm{dep}} = \sigma_p^{-2}\,\mathbf{I}_2$, and the bracket is the Schur complement of equation (4.33). The same coordinates give the covariance of the dependent entries in the limit, $\mathbf{N}_{\mathrm{b}}\boldsymbol{\Sigma}_{\mathrm{red}}\mathbf{N}_{\mathrm{b}}^{\mathsf{T}}$, in the hall the lower block of the limit (4.35): the dependent entries are known through the free ones, each by its row of $\mathbf{N}_{\mathrm{b}}$. The exact balance is this limit, never a substitution made by hand.

### 4.2.9 The fit of the hall state

The fit is run at the tolerance $\sigma_p = 0.01\,\mathrm{K}$ of Table 4.3, and the limit of exact rows is read off equation (4.35).

**The step from the handbook state.** The iteration starts at the forward calculation with the handbook resistance, $\mathbf{z}^{(0)} = (0.160,\ 22.8,\ 27.6)^{\mathsf{T}}$, where both physical rows and the prior row vanish and the reading row is $(22.8 - 23.2)/0.5 = -0.8$, so that $\mathcal{L}(\mathbf{z}^{(0)}) = 0.32$ and $\mathbf{J}^{\mathsf{T}}\mathbf{r}(\mathbf{z}^{(0)}) = (0,\ -1.6,\ 0)^{\mathsf{T}}$: only the reading is out of balance, and it pulls on $T_1$ alone. The normal equations (4.29) are

$$\begin{bmatrix} 45\,001\,600 & -300\,000 & -600\,000 \\ -300\,000 & 10\,004 & 0 \\ -600\,000 & 0 & 10\,000 \end{bmatrix} \Delta\mathbf{z} = \begin{bmatrix} 0 \\ 1.6 \\ 0 \end{bmatrix}, \qquad \Delta\mathbf{z} = \begin{bmatrix} 0.0092296 \\ 0.27694 \\ 0.55378 \end{bmatrix}.$$

Divided by its first entry, the step is $\Delta\mathbf{z} = 0.0092296 \times (1,\ 30.005,\ 60.000)^{\mathsf{T}}$: it lies, to within the tolerance, along the free direction $\mathbf{v}$ of equation (4.30). Table 4.4 lists the iteration. Each line is one iterate: its three entries, the objective, and the largest entry of the gradient (4.28). After the first step the gradient is at round-off and the second step moves no entry by more than $10^{-13}$.

Table 4.4: Gauss–Newton iterations for the hall state from the handbook state (0.160, 22.8, 27.6), with $\sigma_p = 0.01\,\mathrm{K}$.

| Iterate $k$ | $\rho$ [K/kW] | $T_1$ [°C] | $T_2$ [°C] | $\mathcal{L}$ [–] | $\max_i \lvert(\mathbf{J}^{\mathsf{T}}\mathbf{r})_i\rvert$ [kW/K or K$^{-1}$] |
|---|---|---|---|---|---|
| 0 | 0.160000 | 22.8000 | 27.6000 | 0.32000 | 1.6 |
| 1 | 0.169230 | 23.0769 | 28.1538 | 0.09845 | $< 10^{-8}$ |
| 2 | 0.169230 | 23.0769 | 28.1538 | 0.09845 | $< 10^{-8}$ |

**The MAP state.** The fit returns

$$\hat{\mathbf{z}} = \begin{bmatrix} \hat{\rho} \\ \hat{T}_1 \\ \hat{T}_2 \end{bmatrix} = \begin{bmatrix} 0.16923 \\ 23.0769 \\ 28.1538 \end{bmatrix} : \tag{4.49}$$

a resistance of $0.16923\,\mathrm{K/kW}$, a test temperature of 23.077°C and a full-load temperature of 28.154°C, each read straight off its entry, with the effective conductance $\hat{G} = 1/\hat{\rho} = 5.909\,\mathrm{kW/K}$ evaluated at the estimated resistance.

### 4.2.10 The covariance and the resistance

**The covariance.** Inverting the matrix (4.27),

$$\boldsymbol{\Sigma} = \left(\mathbf{J}^{\mathsf{T}}\mathbf{J}\right)^{-1} = \begin{bmatrix} 1.9236 \times 10^{-4} & 5.7685 \times 10^{-3} & 1.1542 \times 10^{-2} \\ 5.7685 \times 10^{-3} & 0.17309 & 0.34611 \\ 1.1542 \times 10^{-2} & 0.34611 & 0.69260 \end{bmatrix} \quad (\sigma_p = 0.01\,\mathrm{K}). \tag{4.50}$$

Its eigenvalues are 0.866 along (nearly) $\mathbf{v}$, and $1.0 \times 10^{-4}$ and $2.2 \times 10^{-8}$ in the two directions the physics fixes, both shrinking as $\sigma_p^2$.

**The entries read off.** The resistance has

$$\hat{\rho} = \hat{z}_1 = 0.16923\,\mathrm{K/kW}, \qquad \sqrt{\Sigma_{11}} = \sqrt{1.9236 \times 10^{-4}} = 0.01387\,\mathrm{K/kW},$$

the test temperature $\sqrt{\Sigma_{22}} = 0.416\,\mathrm{K}$, and the full-load temperature $\hat{T}_2 = \hat{z}_3 = 28.154°\mathrm{C}$ with $\sqrt{\Sigma_{33}} = \sqrt{0.69260} = 0.8322\,\mathrm{K}$. The correlations $\Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$ are 0.99971 between $\rho$ and $T_1$, 0.99993 between $\rho$ and $T_2$, and 0.99964 between $T_1$ and $T_2$.

**The limit of exact rows.** The limit covariance (4.35) gives $\Sigma_{11} = 1/5200$, so that

$$\sqrt{\Sigma_{11}} = 0.013868\,\mathrm{K/kW}, \qquad \hat{\rho} = 880/5200 = 0.169231\,\mathrm{K/kW}, \qquad \hat{T}_2 = 18{+}60\hat{\rho} = 28.1538°\mathrm{C},$$

the resistance being the minimum of equation (4.37); the rest of the chapter carries these limit values, except Section 4.3.8, which says where it does not.

**The objective along the free direction.** Figure 4.3 draws the objective (4.36) on the line $\mathbf{z} = \hat{\mathbf{z}} + t\,\mathbf{v}$ in the limit of exact rows, with $t$ read on the horizontal axis as the resistance $\rho = \hat{z}_1 + t$ in K/kW and the dimensionless objective on the vertical axis: the solid curve is the objective, the dashed curve the prior term and the dotted curve the reading term; the dot marks the minimum and the double arrow joins the two points where the objective has risen by $\frac{1}{2}$. At the minimum $r_3(\hat{\mathbf{z}}) = -0.2462$ and $r_4(\hat{\mathbf{z}}) = 0.3692$, and the slopes cancel, $60 \times (-0.2462) + 40 \times 0.3692 = 0$.

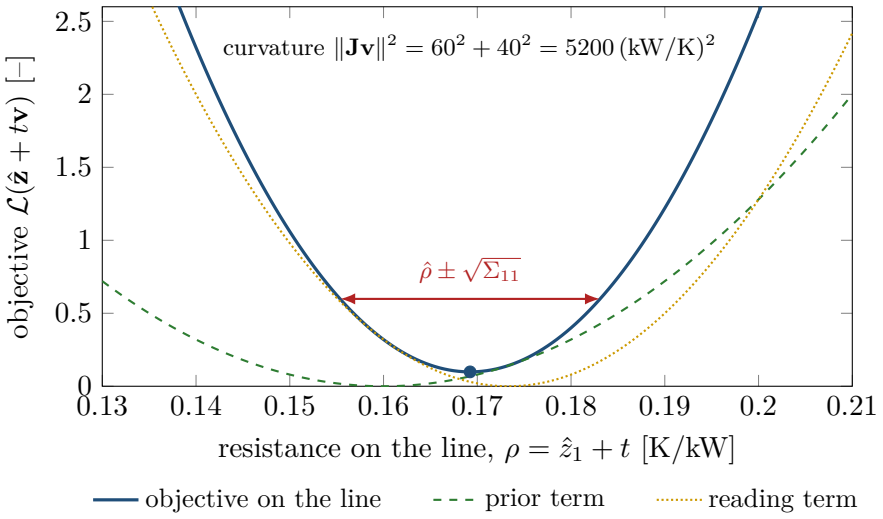



Figure 4.3: The objective on the line $\mathbf{z} = \hat{\mathbf{z}} + t\,\mathbf{v}$ along the free direction $\mathbf{v} = (1, 30, 60)^\mathsf{T}$, in the limit of exact physical rows, with its prior term $\frac{1}{2}(r_4(\hat{\mathbf{z}}) + 40t)^2$ and reading term $\frac{1}{2}(r_3(\hat{\mathbf{z}}) + 60t)^2$; the horizontal axis is the resistance $\hat{z}_1 + t$. Dot: the minimum at $\hat{\rho} = 0.16923\,\mathrm{K/kW}$, $\mathcal{L} = 0.0985$. Double arrow: the points $\hat{\rho} \pm \sqrt{\Sigma_{11}}$, $\sqrt{\Sigma_{11}} = 0.01387\,\mathrm{K/kW}$, where the objective has risen by $\frac{1}{2}$.

**The posterior of the resistance.** By the first marginal of equation (4.31),

$$\rho \mid \mathbf{y} \sim \mathcal{N}(\hat{z}_1,\ \Sigma_{11}) = \mathcal{N}(0.169231,\ 0.013868^2), \tag{4.51}$$

a standard deviation of $0.013868/0.169231 = 0.082$ of the estimate. Figure 4.4 compares it with the prior. The horizontal axis is the resistance and the vertical axis the probability density

in kW/K; the dashed curve is the prior of equation (4.12) and the solid curve the posterior of equation (4.51). The reading has shifted the resistance upward, from 0.160 to 0.169, and narrowed it, from a standard deviation of 0.025 to 0.0139.

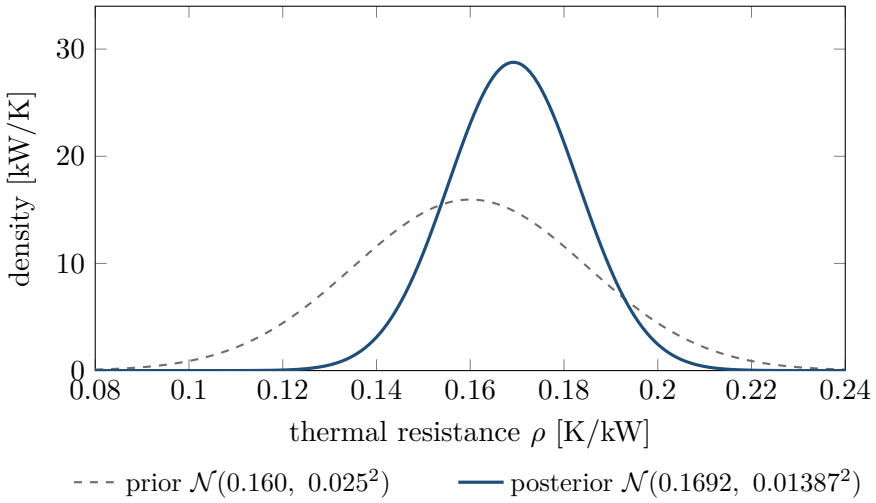



Figure 4.4: Prior and posterior densities of the thermal resistance, the first entry of the state: $\mathcal{N}(0.160, 0.025^2)$ and $\mathcal{N}(0.169231, 0.013868^2)$. The reading moves the resistance up and reduces its spread.

### 4.2.11 The reduced state of the hall

At the tolerance $\sigma_p = 0.01\,\mathrm{K}$ of equation (4.25), $\mathbf{A}_{\mathrm{free}} = (-3000,\ -6000)^{\mathsf{T}}$ and $\mathbf{A}_{\mathrm{dep}} = \mathrm{diag}(100,\ 100)$. From the handbook resistance 0.160 the reduced rows (4.46) are $(-0.8,\ 0)$, $\mathbf{J}_{\mathrm{red}}^{\mathsf{T}}\mathbf{r} = -48$, and one Gauss–Newton step $48/5200 = 0.0092308$ lands on $\hat{\rho} = 0.169231\,\mathrm{K/kW}$, which a second step does not move. The carried entries have $\mathbf{N}_{\mathrm{b}}\boldsymbol{\Sigma}_{\mathrm{red}}\mathbf{N}_{\mathrm{b}}^{\mathsf{T}}$ with diagonal 0.17308 and $0.69231\,\mathrm{K}^2$ and off-diagonal $0.34615\,\mathrm{K}^2$.

Table 4.5 sets the reduced state against the full one. Its first two rows give $\mathbf{N}$ and $\mathbf{J}_{\mathrm{red}}$, which do not depend on the tolerance; the remaining rows give, for the full state at three tolerances and for the reduced state in the last column, the $(\rho, \rho)$ entry of the covariance, its square root, the MAP resistance, and the relative difference of the full entry from $\boldsymbol{\Sigma}_{\mathrm{red}}$. The difference shrinks a hundredfold per decade of $\sigma_p$, as the order $\sigma_p^2$ of the correction in equation (4.48) requires, and at $\sigma_p = 0.001\,\mathrm{K}$ the two agree to every printed digit.

Table 4.5: The reduced state of the hall against the free-entry marginal of the full state. $\mathbf{N}$ is in [–, kW, kW]; the last column is the reduced state, exact in the limit of exact rows.

| Quantity | $\sigma_p = 0.1\,\mathrm{K}$ | $0.01\,\mathrm{K}$ | $0.001\,\mathrm{K}$ | reduced state |
|---|---|---|---|---|
| null-space basis $\mathbf{N}$ | $(1,\ 30,\ 60)^{\mathsf{T}}$ at every tolerance | | | |
| $\mathbf{J}_{\mathrm{red}} = \mathbf{J}_{\mathrm{rows}}\mathbf{N}$ [kW/K] | $(60,\ 40)^{\mathsf{T}}$, information $5200\,(\mathrm{kW/K})^2$ | | | |
| $\Sigma_{11}$, $\boldsymbol{\Sigma}_{\mathrm{red}}$ $[(\mathrm{K/kW})^2]$ | $1.97568 \times 10^{-4}$ | $1.92361 \times 10^{-4}$ | $1.92308 \times 10^{-4}$ | $1.92308 \times 10^{-4}$ |
| standard deviation [K/kW] | 0.014056 | 0.013869 | 0.013868 | 0.013868 |
| $\hat{\rho}$ [K/kW] | 0.169119 | 0.169230 | 0.169231 | 0.169231 |
| relative difference from $\boldsymbol{\Sigma}_{\mathrm{red}}$ [–] | $2.7 \times 10^{-2}$ | $2.8 \times 10^{-4}$ | $2.8 \times 10^{-6}$ | – |

## 4.3 The answers (step 4)

Each question of Table 4.2 is read from the posterior of Section 4.2.

### 4.3.1 Uncertainty propagation and the delta method

The rows of the hall are linear, so the entries of $\boldsymbol{\Sigma}$ are exact variances. In most plant models an output is a nonlinear function of the state, $f(\mathbf{z})$, and the posterior is Gaussian at best approximately, $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$.

**Linearization of the output.** The first-order Taylor expansion of $f$ about $\hat{\mathbf{z}}$ is

$$f(\mathbf{z}) \approx f(\hat{\mathbf{z}}) + \nabla f(\hat{\mathbf{z}})^{\mathsf{T}}(\mathbf{z} - \hat{\mathbf{z}}).$$

Writing $\mathbf{g} = \nabla f(\hat{\mathbf{z}})$ for the column gradient,

$$f(\mathbf{z}) \approx f(\hat{\mathbf{z}}) + \mathbf{g}^{\mathsf{T}}(\mathbf{z} - \hat{\mathbf{z}}). \tag{4.52}$$

Taking the posterior expectation, with $\mathbb{E}[\mathbf{z} - \hat{\mathbf{z}} \mid \mathbf{y}] = \mathbf{0}$ because $\hat{\mathbf{z}}$ is the mean of the Gaussian,

$$\mu_f = \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] \approx f(\hat{\mathbf{z}}).$$

Taking the variance, the constant $f(\hat{\mathbf{z}})$ drops out, and the variance of the linear combination $\mathbf{g}^{\mathsf{T}}(\mathbf{z} - \hat{\mathbf{z}}) = \sum_i g_i(z_i - \hat{z}_i)$ is the double sum of its covariances,

$$\operatorname{Var}(f(\mathbf{z}) \mid \mathbf{y}) \approx \operatorname{Var}\Big(\sum_i g_i z_i \;\Big|\; \mathbf{y}\Big) = \sum_i \sum_j g_i\, g_j\, \operatorname{Cov}(z_i, z_j \mid \mathbf{y}),$$

which in matrix form is

$$\operatorname{Var}(f(\mathbf{z}) \mid \mathbf{y}) \approx \mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}. \tag{4.53}$$

This is the first-order delta method, one of the most reused formulas in probabilistic plant modeling. Its only approximation is the linearization (4.52); when $f$ is linear and the posterior Gaussian, it is exact.

**Application to the thermal system.** The full-load temperature is the third entry of the state, so its gradient is the third unit vector and the quadratic form picks one diagonal entry,

$$\mathbf{g} = \frac{\partial T_2}{\partial \mathbf{z}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g} = \Sigma_{33}.$$

The same temperature can also be predicted as a derived quantity of the resistance alone, $f(\mathbf{z}) = T_c + Q_2 \rho = 18 + 60\rho$, with

$$\mathbf{g} = \frac{\partial f}{\partial \mathbf{z}} = \begin{bmatrix} 60 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g} = 60^2\, \Sigma_{11}.$$

In the limit the two are equal, since the limit covariance (4.35) has $\Sigma_{33} = 3600/5200 = 60^2\, \Sigma_{11}$; at a finite tolerance they differ by $\sigma_p^2$, equation (4.34), because the entry carries the discrepancy of the full-load balance and the prediction from $\rho$ does not. The two calculations, reading the variance off the covariance and propagating the resistance's uncertainty through a gradient, agree in the limit because they are the same transformation of the same uncertainty: an output that is an entry of the state needs no propagation, and an output that is not needs its gradient.

### 4.3.2 The gradient and the variance budget

**Physical meaning of the gradient.** The first entry of the gradient $\mathbf{g} = (60,\ 0,\ 0)^{\mathsf{T}}$ of the prediction $18 + 60\rho$, 60 kW, means that a change of 0.01 K/kW in the resistance changes the full-load temperature by $\Delta T_2 = 60 \times 0.01 = 0.6\,\mathrm{K}$. The gradient is the link between physical uncertainty and operating consequence. A state entry can have a large posterior variance and still contribute little to an output that is insensitive to it, with $g_i \approx 0$; conversely, a modest uncertainty in a highly influential entry can dominate the uncertainty of a prediction.

**Several uncertain entries.** Suppose the supply-air temperature is uncertain as well, the controller holding it only near its set point, so that the full-load temperature $T = T_c + Q\rho$ depends on the sub-vector $\boldsymbol{\theta} = (T_c, \rho)^{\mathsf{T}}$ of the state, with posterior covariance $\boldsymbol{\Sigma}_\theta$. The gradient is $\mathbf{g} = (\partial T/\partial T_c,\ \partial T/\partial \rho)^{\mathsf{T}} = (1,\ Q)^{\mathsf{T}}$, and equation (4.53) gives

$$\mathrm{Var}(T) = \mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma}_\theta\, \mathbf{g} = \begin{bmatrix} 1 & Q \end{bmatrix} \begin{bmatrix} \mathrm{Var}(T_c) & \mathrm{Cov}(T_c, \rho) \\ \mathrm{Cov}(T_c, \rho) & \mathrm{Var}(\rho) \end{bmatrix} \begin{bmatrix} 1 \\ Q \end{bmatrix}.$$

Multiplying out,

$$\mathrm{Var}(T) = \mathrm{Var}(T_c) + Q^2\ \mathrm{Var}(\rho) + 2Q\ \mathrm{Cov}(T_c, \rho). \tag{4.54}$$

This is the variance budget of the output: one term per entry, each its sensitivity squared times its variance, plus a cross term from their covariance. The first term is the spread the supply air alone puts on the temperature; the second, the spread the resistance alone puts on it, $Q^2\ \mathrm{Var}(\rho)$; the third corrects for the two moving together in the posterior. If the readings cannot separate the two entries, the posterior leaves them correlated, and the cross term either adds to the diagonal terms or cancels part of them. With a negative correlation and same-signed sensitivities it is negative, and the output can be well known although each entry is poorly known: the readings then fix the combination $T_c + Q\rho$ that the output depends on, not the entries separately. Adding the two individual variances without the cross term is wrong in either direction, which is why the joint posterior, not its marginals, must be kept; Exercise 4.2 works a case in which ignoring the cross term overstates the spread by more than half.

### 4.3.3 The density of a transformed quantity

The delta method gives a derived quantity a mean and a variance. Some questions need more: the whole density of the quantity, or a tail probability far from the center, where the shape matters and a Gaussian with the right variance can still be wrong. The tool for these is the *change of variables*, and it is used throughout the book whenever a quantity is computed in one coordinate and reported in another.

**Probability is conserved; density is not.** Let $x$ have the density $p_x(x)$, and let $u = h(x)$ be a monotone function with inverse $x = h^{-1}(u)$. A short interval $[x, x + \mathrm{d}x]$ is mapped onto a short interval $[u, u + \mathrm{d}u]$, and the two carry the same probability, because they hold the same outcomes:

$$p_u(u)\,|\mathrm{d}u| = p_x(x)\,|\mathrm{d}x|.$$

Dividing by $|\mathrm{d}u|$ gives the density of the transformed quantity,

$$p_u(u) = p_x\big(h^{-1}(u)\big)\,\left|\frac{\mathrm{d}x}{\mathrm{d}u}\right|. \tag{4.55}$$

The first factor evaluates the old density at the point that maps to $u$. The second, the *change-of-variable factor*, records how much the map stretches or squeezes the axis there: where a small change of $u$ corresponds to a large change of $x$, the probability of the long $x$-interval is packed

into a short $u$-interval, and the density of $u$ must be larger to hold it. Without the factor, $p_x(h^{-1}(u))$ is not a density of $u$ at all: it does not in general integrate to one. A monotone map also carries quantiles over unchanged: if $h$ is increasing, $\Pr(u \le h(x^*)) = \Pr(x \le x^*)$, so the median of $u$ is $h$ of the median of $x$, while the mode, which depends on the factor, moves.

**The conductance.** For the conductance $G = 1/\rho$ of the hall, the inverse map is $\rho = 1/G$ and $|\mathrm{d}\rho/\mathrm{d}G| = 1/G^2$, so equation (4.55) gives

$$p_G(G) = \frac{1}{G^2}\, p(\rho = 1/G \mid \mathbf{y}).$$

It is skewed to the right of $\hat{G}$ because the factor $1/G^2$ is larger at small $G$, which squeezes the left of the curve and stretches its right.

**Two routes to a derived quantity.** The delta method and the change of variables answer the same question with different approximations. For a scalar entry with posterior $x \mid \mathbf{y} \approx \mathcal{N}(\hat{x}, \sigma^2)$ and a monotone output $u = h(x)$:

- the *delta method* linearizes $h$ at $\hat{x}$, with gradient $g = \mathrm{d}u/\mathrm{d}x$ evaluated there, and reports $u \mid \mathbf{y} \approx \mathcal{N}(h(\hat{x}),\ g^2\sigma^2)$, a Gaussian in $u$;
- the *change of variables* keeps $h$ exact and reports $p_u(u) = p_x(h^{-1}(u))\,|\mathrm{d}x/\mathrm{d}u|$, with the Gaussian of $x$ as its only approximation.

The two factors are reciprocals: the delta method's gradient $\mathrm{d}u/\mathrm{d}x$ is the inverse of the change-of-variable factor $\mathrm{d}x/\mathrm{d}u$, but the delta method evaluates it once, at $\hat{x}$, while the change of variables evaluates it at every $u$. The delta method is the change of variables with the map replaced by its tangent, which is why the two agree on the spread near the center and differ in the tails. When $h$ is linear the factor is constant and the two coincide. When $h$ bends across the posterior, the delta method puts a symmetric Gaussian on a quantity whose density is skewed, and can put probability where the quantity cannot go: below zero for a conductance or a rate, outside $(0, 1)$ for a probability such as a probability of detection.

> **Principle 4.1.** *The gradient carries a spread; the change-of-variable factor carries a density. Report $\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}$ when a standard deviation is the answer, and transform the density when the answer is a shape, a tail probability or a bound near the edge of what the quantity can be.*

### 4.3.4 Credible intervals

A commissioning record often needs an interval for an uncertain quantity as well as its mean. For the Gaussian marginal $T_2 \mid \mathbf{y} \sim \mathcal{N}(\hat{z}_3, \Sigma_{33})$ of equation (4.31) the central 95 % credible interval is the interval that leaves 2.5 % of the posterior in each tail. Since $\Phi(1.960) = 0.975$,

$$\left[\hat{z}_3 - 1.960\,\sqrt{\Sigma_{33}},\ \hat{z}_3 + 1.960\,\sqrt{\Sigma_{33}}\right]. \tag{4.56}$$

The same form serves any entry, with $\hat{z}_i$ and $\sqrt{\Sigma_{ii}}$, and a derived quantity, with $f(\hat{\mathbf{z}})$ and $\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}$ from the delta method (4.53).

The interval and a threshold probability carry different information. The interval describes the range that holds most of the posterior probability; a threshold probability answers directly whether the plant meets a stated limit. For a release decision the threshold probability is usually the more actionable of the two. Neither is meaningful apart from the model assumptions, the quality of the measurement and the sources of uncertainty that the model includes; Section 4.6.6 lists those it leaves out.

### 4.3.5 The temperature-threshold probability

The question of equation (4.10) needs one marginal of the posterior. The marginal of the full-load temperature is read off the third entry of the state, equation (4.31): $T_2 \mid \mathbf{y} \sim \mathcal{N}(\hat{z}_3,\ \Sigma_{33})$.

**The standardized temperature.** For a Gaussian variable the standardized variable $(T_2 - \hat{z}_3)/\sqrt{\Sigma_{33}}$ has the standard normal distribution, and subtracting a constant and dividing by a positive constant preserve the inequality, so

$$\Pr(T_2 < T_{\max} \mid \mathbf{y}) = \Pr\Big(\frac{T_2 - \hat{z}_3}{\sqrt{\Sigma_{33}}} < \frac{T_{\max} - \hat{z}_3}{\sqrt{\Sigma_{33}}} \;\Big|\; \mathbf{y}\Big) = \Phi\Big(\frac{T_{\max} - \hat{z}_3}{\sqrt{\Sigma_{33}}}\Big). \tag{4.57}$$

The probability of exceeding the limit is its complement, $1 - \Pr(T_2 < T_{\max} \mid \mathbf{y})$.

**The prior marginal.** Before the test reading the state had the prior formed by the physical and prior rows alone. Its Jacobian $\mathbf{J}_0$ is $\mathbf{J}$ of equation (4.25) without the reading's line, its mean is the handbook state $\mathbf{z}^{(0)} = (0.160,\ 22.8,\ 27.6)^{\mathsf{T}}$, and its covariance $\mathbf{\Sigma}_0 = (\mathbf{J}_0^{\mathsf{T}}\mathbf{J}_0)^{-1}$ tends, by the argument of equation (4.35) with $\mathbf{J}_0\mathbf{v} = (0,\ 0,\ 40)^{\mathsf{T}}$ in place of $\mathbf{J}\mathbf{v}$, to $\mathbf{v}\mathbf{v}^{\mathsf{T}}/40^2 = \sigma_0^2\,\mathbf{v}\mathbf{v}^{\mathsf{T}}$. Its third diagonal entry is $(\mathbf{\Sigma}_0)_{33} = 60^2 \times 0.025^2 = 2.25\,\mathrm{K}^2$, so the full-load temperature had the prior marginal $\mathcal{N}(27.6,\ 1.5^2)$, the third entry of $\mathbf{z}^{(0)}$ with $\sqrt{(\mathbf{\Sigma}_0)_{33}} = 1.5\,\mathrm{K}$.

**The question through the resistance.** In the limit of exact rows the posterior lies on the line $\mathbf{z} = \hat{\mathbf{z}} + t\,\mathbf{v}$ through the MAP state along the free direction, $\mathbf{\Sigma} = \mathbf{v}\mathbf{v}^{\mathsf{T}}/5200$ by equation (4.35). On that line the third entry is $\hat{z}_3 + 60t$ and the first $\hat{z}_1 + t$, so the event $\{T_2 < T_{\max}\}$ is the event $\{t < t^*\}$ with $t^* = (T_{\max} - \hat{z}_3)/60$, which is the event that the resistance stays below

$$\rho^* = \hat{z}_1 + t^* = \hat{z}_1 + \frac{T_{\max} - \hat{z}_3}{60}. \tag{4.58}$$

With the marginal (4.31) of the first entry,

$$\Pr(\rho < \rho^* \mid \mathbf{y}) = \Phi\Big(\frac{\rho^* - \hat{z}_1}{\sqrt{\Sigma_{11}}}\Big) = \Phi\Big(\frac{t^*}{\sqrt{\Sigma_{11}}}\Big),$$

which is the probability (4.57): both standardized distances are $t^*$ divided by the spread of $t$, since the third entry moves 60 times as far as the first along $\mathbf{v}$ and $\sqrt{\Sigma_{33}} = 60\sqrt{\Sigma_{11}}$ in the limit. Reading the third entry and propagating the first one along the free direction are the same calculation; the full-load question is a question about the resistance.

**The dependence on the value read.** The information matrix (4.27), and with it $\mathbf{\Sigma}$, does not contain the value observed, so a different reading leaves every standard deviation where it is and moves only the mean, along one vector. The reading enters the normal equations only through the part $-y_T/\sigma_T$ of the third entry of the residual vector that depends on $y_T$; that row's line of the Jacobian is $\mathbf{h}^{\mathsf{T}}/\sigma_T$, so differentiating $\hat{\mathbf{z}} = \mathbf{z} - \mathbf{\Sigma}\mathbf{J}^{\mathsf{T}}\mathbf{r}(\mathbf{z})$ with respect to $y_T$ gives the gain vector

$$\frac{\partial \hat{\mathbf{z}}}{\partial y_T} = \frac{\mathbf{\Sigma}\mathbf{h}}{\sigma_T^2}, \tag{4.59}$$

the second column of $\mathbf{\Sigma}$ divided by $\sigma_T^2$. The threshold probability (4.57) is computed with a spread fixed before the reading and a mean set by its value, so the reading can move it in either direction.

### 4.3.6 The posterior state and the full-load temperature

**General form.** The three estimation questions, the resistance of the air path and the hall temperature at the test and at full load, share one form: the posterior of a quantity $f(\mathbf{z})$, equation (2.9), summarized by its moments (2.10) and its central interval (2.11).

**Specialized form.** Equations (2.10) and (2.11) with $f(\mathbf{z}) = z_i$, the entry asked about, and $p(\mathbf{z} \mid \mathbf{y})$ the Gaussian posterior of step 3: the mean is $\hat{z}_i$, the variance $\Sigma_{ii}$, by the marginals (4.31), and the interval is equation (4.56). The rows are linear and Gaussian, so these are the exact moments, with no approximation to set beside them.

**Building blocks.** The MAP state and the limit covariance (4.35), with its limit values of Section 4.2.10; the prior marginal $\mathcal{N}(27.6,\ 1.5^2)$ of Section 4.3.5, for the prior interval.

**The posterior state.** Table 4.6 collects it: each row is one entry of $\hat{\mathbf{z}}$, with its posterior mean and its standard deviation $\sqrt{\Sigma_{ii}}$, in the limit $\sigma_p \to 0$; the full-load spread is twice the test spread, the ratio 60/30 of the entries of $\mathbf{v}$.

Table 4.6: The posterior state of the equipment hall after the test reading, in the limit of exact physical rows.

| State entry | Posterior mean | Standard deviation |
|---|---|---|
| $\rho$ [K/kW] | 0.16923 | 0.01387 |
| $T_1$ [°C; sd in K] | 23.077 | 0.416 |
| $T_2$ [°C; sd in K] | 28.154 | 0.832 |

**The diagonal of the limit covariance.** The limit covariance (4.35) gives $\Sigma_{11} = 1/5200 = 1.9231 \times 10^{-4}\,(\mathrm{K/kW})^2$, $\Sigma_{22} = 900/5200 = 0.17308\,\mathrm{K}^2$ ($\sqrt{\Sigma_{22}} = 0.4160\,\mathrm{K}$) and

$$\Sigma_{33} = \mathrm{Var}(T_2 \mid \mathbf{y}) = \frac{3600}{5200} = 0.69231\,\mathrm{K}^2, \qquad \sqrt{\Sigma_{33}} = 0.8321\,\mathrm{K}. \tag{4.60}$$

**The credible interval.** By equation (4.56), $28.1538 \pm 1.960 \times 0.8321 = 28.1538 \pm 1.6308$, so that

$$T_2 \in [26.52,\ 29.78]^\circ\mathrm{C}$$

with 95 % posterior probability under the stated model; the prior interval was $[24.66,\ 30.54]^\circ$C, 1.8 times as wide.

**The conductance beside the resistance.** The conductance is a derived quantity, $G(\mathbf{z}) = 1/z_1$, so the delta method (4.53) reads it from the same posterior, with

$$\mathbf{g} = \frac{\partial G}{\partial \mathbf{z}} = \begin{bmatrix} -1/\hat{z}_1^2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -34.92 \\ 0 \\ 0 \end{bmatrix} (\mathrm{kW/K})^2, \qquad \mathrm{Var}(G \mid \mathbf{y}) \approx \mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g} = \frac{\Sigma_{11}}{\hat{z}_1^4},$$

which gives $\hat{G} = 1/\hat{\rho} = 5.909\,\mathrm{kW/K}$ with $\mathrm{sd}(G \mid \mathbf{y}) \approx 0.484\,\mathrm{kW/K}$ and the 5–95 % band $[5.113,\ 6.706]\,\mathrm{kW/K}$. The exact band, by the change of variables (4.55), is $[5.207,\ 6.830]\,\mathrm{kW/K}$, higher by 0.095 and $0.124\,\mathrm{kW/K}$; the median is $5.909\,\mathrm{kW/K}$ on both routes.

### 4.3.7 The threshold probability at full load

**General form.** A requirement that a quantity stay below a limit is the posterior probability of the event, equation (2.12), compared with an assurance level $\kappa$.

**Specialized form.** Equation (2.12) with $f(\mathbf{z}) = -T_2 = -z_3$, $f^* = -T_{\max}$ and $p(\mathbf{z} \mid \mathbf{y})$ the Gaussian posterior of step 3 becomes the normal tail (4.57) of the third entry. The commissioning test states no assurance level $\kappa$; the probability is reported without an accept verdict. The posterior is exactly Gaussian, so the tail is the exact probability under the model.

**Building blocks.** The third entry $\hat{z}_3$ and the variance $\Sigma_{33}$ of equation (4.60), and the standard normal distribution function $\Phi$.

**The probability.** With $\hat{z}_3 = 28.1538\,^\circ\mathrm{C}$ and $\sqrt{\Sigma_{33}} = 0.8321\,\mathrm{K}$, the standardized distance in equation (4.57) is $(28 - 28.1538)/0.8321 = -0.1849$, and

$$\Pr(T_2 < 28\,^\circ\mathrm{C} \mid \mathbf{y}) = \Phi(-0.1849) = 0.4267. \tag{4.61}$$

The probability of exceeding the limit is $1 - 0.4267 = 0.5733$. The fit at $\sigma_p = 0.01\,\mathrm{K}$ gives 0.4267 as well, with $\sqrt{\Sigma_{33}} = 0.83223\,\mathrm{K}$ against the limit's $0.83205\,\mathrm{K}$.

**Against the prior.** Under the prior marginal $\mathcal{N}(27.6,\ 1.5^2)$ of Section 4.3.5, before the test reading, the probability that the full-load temperature stays below the limit was

$$\Pr(T_2 < 28) = \Phi\left(\frac{28 - 27.6}{1.5}\right) = \Phi(0.2667) = 0.6051.$$

Figure 4.5 draws the answer. The horizontal axis is the full-load temperature and the vertical axis its density in $\mathrm{K}^{-1}$. The solid curve is the posterior marginal, the dashed curve the prior marginal, and the vertical line the limit $T_{\max}$. The shaded area under the posterior to the left of the line is the probability (4.61); the unshaded area under it to the right is the probability of exceeding the limit. The posterior is narrower than the prior and centered to the right of it, across the limit.

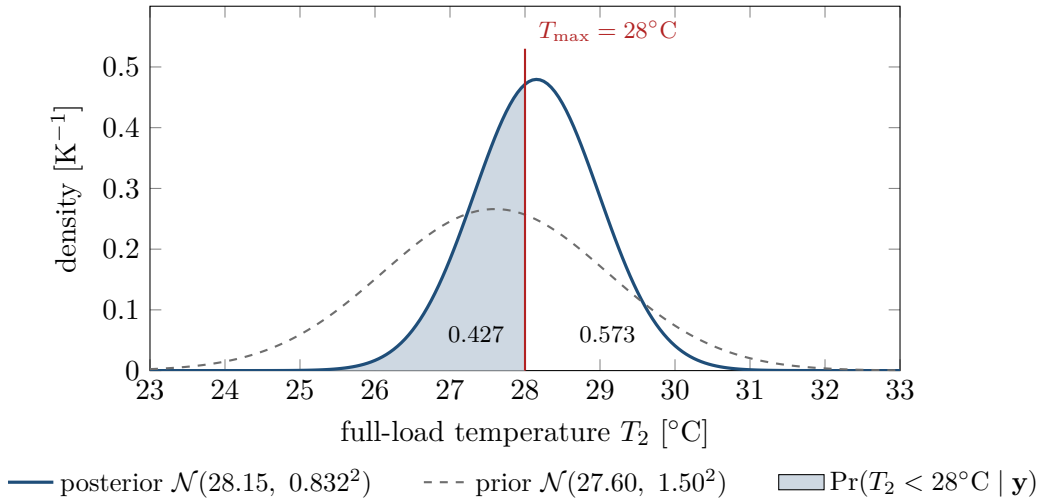



Figure 4.5: The threshold probability of equation (4.57). Posterior marginal of the full-load temperature, $\mathcal{N}(28.1538, 0.8321^2)$, with the area below $T_{\max} = 28\,^\circ\mathrm{C}$ shaded ($\Phi(-0.1849) = 0.4267$), and the prior marginal $\mathcal{N}(27.6, 1.5^2)$, under which the same probability is $\Phi(0.2667) = 0.6051$.

### 4.3.8 The set point at a scenario node

The commissioning test states no assurance level, but the release can ask what a level it chooses would require of the handle the loop provides. The question of the last row of Table 4.2 takes $\kappa = 0.95$.

**General form.** The threshold form (2.12) applies to a quantity $f_r(\mathbf{z})$ that depends on an input $r$ pinned at a scenario node, and the question asks for the largest input that meets the level,

$$r_\kappa = \max\{\, r : \Pr(f_r(\mathbf{z}) \geq f^* \mid \mathbf{y}) \geq \kappa \,\}. \tag{4.62}$$

**Specialized form.** Equations (2.12) and (4.62) with $f_r(\mathbf{z}) = -T_2^*(r)$, $f^* = -T_{\max}$ and $p(\mathbf{z} \mid \mathbf{y})$ the Gaussian posterior of step 3. The scenario node pins the set point at $r$ and re-solves the full-load balance (4.4) there,

$$T_2^*(r) = r + Q_2\rho, \qquad \mathbf{g} = \frac{\partial T_2^*}{\partial \mathbf{z}} = \begin{bmatrix} Q_2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 60 \\ 0 \\ 0 \end{bmatrix}. \tag{4.63}$$

The what-if is closed loop in the form this part of the book allows: the set point is pinned, the discharge controller is taken to reach it, and no actuator appears in the graph. This is what lets $r$ enter as a constant, and it fails if the unit cannot reach $r$ at 60 kW, and it fails if the discharge sensor misreads, since the air then leaves at other than $r$. The test row stays at 18°C, where the test was run, and the node shares its resistance: this is what carries the reading to the node, and it fails if the path changes with the supply temperature (Section 4.6.6). The gradient does not contain $r$. With the node's balance held to the same tolerance as the full-load row, the moments follow from the state's and from equation (4.34),

$$\mathbb{E}[T_2^*(r) \mid \mathbf{y}] = r + Q_2\hat{z}_1, \qquad \mathrm{Var}(T_2^*(r) \mid \mathbf{y}) = \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} + \sigma_p^2 = \Sigma_{33}. \tag{4.64}$$

The mean moves one for one with $r$ and the spread does not move at all: a pinned input that enters additively leaves the spread alone; one that multiplies an uncertain entry, as $Q$ does, scales it, and the spread of $T$ pinned at 30, 45 and 60 kW is 0.416, 0.624 and 0.832 K. The normal tail (4.57) of this marginal is

$$\Pr(T_2^*(r) < T_{\max} \mid \mathbf{y}) = \Phi\!\left(\frac{T_{\max} - r - Q_2\hat{z}_1}{\sqrt{\Sigma_{33}}}\right), \tag{4.65}$$

which rises as $r$ falls, so the largest set point of equation (4.62) is the one at which it equals $\kappa$. Inverting the normal distribution function,

$$r_\kappa = T_{\max} - Q_2\hat{z}_1 - \Phi^{-1}(\kappa)\,\sqrt{\Sigma_{33}}. \tag{4.66}$$

The last term, $\Phi^{-1}(\kappa)\sqrt{\Sigma_{33}}$, is the *back-off*: the distance the set point stands inside the value at which the mean alone would sit on the limit, because the temperature is uncertain. The name is fixed here for this one node; its general form belongs to Chapter 6.

**Building blocks.** The entries $\hat{z}_1$ and $\Sigma_{33}$ of the fit, equations (4.49) and (4.50), and $\Phi$ with its inverse. This subsection carries the fit at $\sigma_p = 0.01$ K, not the limit: the spread is 0.8322 K at that tolerance and 0.8321 K in the limit of Section 4.2.10. The rows are linear, so each is exact on its own route, and the difference is the tolerance's.

**The set point at** 0.95**.** Table 4.7 gives the node at five set points: the first row is the set point, the second the mean of equation (4.64), and the third the probability (4.65). Its first column reproduces the answer (4.61).

With $\Phi^{-1}(0.95) = 1.6449$, equation (4.66) gives

$$r_{95} = 28 - 10.1538 - 1.6449 \times 0.8322 = 16.4773°\mathrm{C}, \tag{4.67}$$

Table 4.7: The full-load temperature at the scenario node $T_2^*(r) = r + Q_2\rho$, with $\sqrt{\Sigma_{33}} = 0.8322\,\mathrm{K}$ at every set point ($\sigma_p = 0.01\,\mathrm{K}$).

| Set point $r$ [°C] | 18.0 | 17.5 | 17.0 | 16.5 | 16.0 |
|---|---|---|---|---|---|
| Mean of $T_2^*$ [°C] | 28.1538 | 27.6538 | 27.1538 | 26.6538 | 26.1538 |
| $\Pr(T_2^* < 28°\mathrm{C})$ [–] | 0.4267 | 0.6613 | 0.8454 | 0.9471 | 0.9867 |

of which the back-off is $1.6449 \times 0.8322 = 1.3689\,\mathrm{K}$. Panel (a) of Figure 4.6 draws the two nodes. Its horizontal axis is the temperature at the node and its vertical axis the density; the dashed curve is the node at $r = 18°\mathrm{C}$, the solid curve the node at $r_{95}$, and the vertical line the limit. The curves have one shape, moved along the axis, and the area under the solid curve to the left of the line is 0.95.

**A second reading.** A further reading at the test condition, its error independent of the first, changes this answer in two ways, and they are known at different times. Its row $(T_1 - y_{T,2})/\sigma_T$ has the line $\mathbf{j} = (0,\ 2,\ 0)$, and information adds, equation (3.27), so

$$\boldsymbol{\Sigma}_2 = \left(\mathbf{J}^{\mathsf{T}}\mathbf{J} + \mathbf{j}^{\mathsf{T}}\mathbf{j}\right)^{-1} \tag{4.68}$$

does not contain the value read. The spread at the node falls from 0.8322 to $\sqrt{(\boldsymbol{\Sigma}_2)_{33}} = 0.6399\,\mathrm{K}$ whatever the reading returns, which releases $1.6449 \times (0.8322 - 0.6399) = 0.3164\,\mathrm{K}$ of back-off for certain (computed before rounding). The value read moves the mean instead. The old rows are balanced at $\hat{\mathbf{z}}$, so the gradient of the enlarged objective there comes from the new row alone, and one step of the normal equations (4.29) gives $\hat{\mathbf{z}} + \boldsymbol{\Sigma}_2\mathbf{h}(y_{T,2} - \mathbf{h}^{\mathsf{T}}\hat{\mathbf{z}})/\sigma_T^2$, the gain (4.59) with $\boldsymbol{\Sigma}_2$ in place of $\boldsymbol{\Sigma}$. Through the gradient of equation (4.63) the mean of $Q_2\rho$ moves by $c = \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_2\mathbf{h}/\sigma_T^2 = 0.81806$ per kelvin of innovation $y_{T,2} - \mathbf{h}^{\mathsf{T}}\hat{\mathbf{z}}$. Equation (4.66) after the reading therefore differs from its value before by

$$\Delta r_\kappa(y_{T,2}) = \Phi^{-1}(\kappa)\Big(\sqrt{\Sigma_{33}} - \sqrt{(\boldsymbol{\Sigma}_2)_{33}}\Big) - c\,(y_{T,2} - \mathbf{h}^{\mathsf{T}}\hat{\mathbf{z}}), \tag{4.69}$$

a sure release and a term that takes either sign. Table 4.8 evaluates it. Each column is a value of the second reading, the third being the predicted value $\mathbf{h}^{\mathsf{T}}\hat{\mathbf{z}}$; the rows give $r_{95}$ after the reading and its change from equation (4.67), computed before rounding.

Table 4.8: The largest set point at 0.95 after a second reading at the test condition, from $r_{95} = 16.4773°\mathrm{C}$ before it. The spread after the reading is 0.6399 K in every column ($\sigma_p = 0.01\,\mathrm{K}$).

| Reading $y_{T,2}$ [°C] | 22.5 | 23.0 | 23.0769 | 23.4 | 23.6 | 24.0 |
|---|---|---|---|---|---|---|
| $r_{95}$ after [°C] | 17.2657 | 16.8566 | 16.7937 | 16.5294 | 16.3658 | 16.0386 |
| Change [K] | +0.7883 | +0.3793 | +0.3164 | +0.0521 | −0.1115 | −0.4388 |

The last two columns are the surprise: the band narrowed, and the allowable set point fell. Equation (4.69) vanishes at

$$y_{T,2}^{\dagger} = \mathbf{h}^{\mathsf{T}}\hat{\mathbf{z}} + \frac{0.3164}{c} = 23.4637°\mathrm{C},$$

and any reading above it lowers $r_{95}$. How often that happens follows from the predictive (2.13) of the reading, which for $Y = \mathbf{h}^{\mathsf{T}}\mathbf{z} + \epsilon$ under the Gaussian posterior is $\mathcal{N}(\mathbf{h}^{\mathsf{T}}\hat{\mathbf{z}},\ \mathbf{h}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{h} + \sigma_T^2) = \mathcal{N}(23.0769,\ 0.6505^2)$ at this tolerance. Under it the probability of a reading above $y_{T,2}^{\dagger}$ is 0.2761.

The innovation has mean zero and equation (4.69) is linear in it, so the expected change is the sure release, 0.3164 K, and the expected $r_{95}$ after the reading is 16.7937°C.

For $\kappa$ above one half, averaged over the predictive, a tighter posterior raises the set point; the limit does not enter. With $s$ and $s_2$ the spreads before and after, the expected change is $\Phi^{-1}(\kappa)(s - s_2)$: zero at $\kappa = 0.5$, and $-0.1009$ K at $\kappa = 0.3$ (limit values $s = 0.83205$ and $s_2 = 0.63960$ K). The move of the mean has variance $s^2 - s_2^2$, so the share of readings that lower the set point is $\Phi\big(-\Phi^{-1}(\kappa)\sqrt{(s - s_2)/(s + s_2)}\big)$, also free of the limit. The gain holds on average, not reading by reading, and the average is over the predictive, not over repeats at this hall: given the true $\rho$ the innovation has mean $Q_1(\rho - \hat{z}_1)$, and the expected change is negative above $\rho = 0.18212$ K/kW, 0.93 standard deviations above the mean, a value of posterior probability 0.1763. A cap at 18°C would trim the expectation by 0.0021 K. Chapter 16 carries this to the two-unit hall.

Panel (b) of Figure 4.6 draws it. In the upper part the horizontal axis is the second reading and the vertical axis $r_{95}$ after it: the solid line is equation (4.69) added to $r_{95}$, the dots are the columns of Table 4.8, and the dashed line is the level before the reading. The lower part, on the same horizontal axis, is the predictive density of the reading. The dotted vertical line marks $y^{\dagger}_{T,2}$, where the solid line crosses the dashed one, and the shaded area to its right is the share of readings that lower the set point. The solid line stands above the dashed one at the center of the density by the sure release.

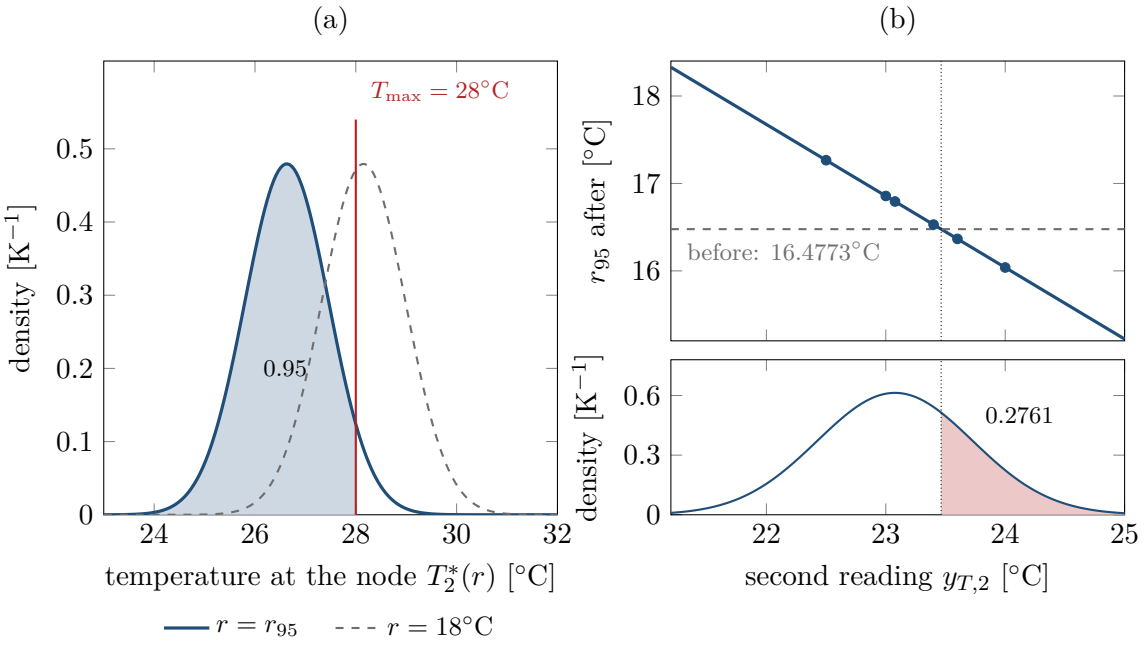



Figure 4.6: The set point at a scenario node ($\sigma_p = 0.01$ K). (a) The normal of $T_2^*(r)$, with mean $r + 10.1538$°C and standard deviation 0.8322 K, at $r = 18$°C and at $r_{95} = 16.4773$°C, with the limit; the shaded area is 0.95. (b) Upper part: $r_{95}$ after a second reading, $16.7937 - 0.81806\,(y_{T,2} - 23.0769)$°C, against that reading, with the level before it dashed. Lower part: the predictive density $\mathcal{N}(23.0769,\ 0.6505^2)$ of the reading. Dotted: the reading 23.4637°C; the shaded area above it, 0.2761, is the share of readings that lower the set point.

**The draws.** A check that does not use equation (4.69) draws the state from the posterior and the sensor error from its model, forms the reading, and refits by the matrix for each draw: the mean after it is $\boldsymbol{\Sigma}_2(\mathbf{J}^{\mathsf{T}}\mathbf{J}\,\hat{\mathbf{z}} + \mathbf{h}\,y_{T,2}/\sigma_T^2)$, and the spread is $\sqrt{(\boldsymbol{\Sigma}_2)_{33}} = 0.6399$ K. Over $N = 10^6$ draws (seed 4, one run) the mean change is 0.3156 K (standard error 0.0005) against the closed

form's 0.3164 K, a gap of $-1.46$ standard errors, and the share lowered is 0.2765 (standard error 0.0004) against the closed form's 0.2761, a gap of $+0.91$ standard errors. The run does not separate either from the closed form, and the closed form is the reported value.

## 4.4 The checks (step 5)

Obtaining a posterior is not the same as establishing that it is credible.

### 4.4.1 Algebraic and physical verification

**Algebraic verification.** At the fitted state of equation (4.49), with $\sigma_p = 0.01\,\mathrm{K}$, the residual vector (4.23) is $\mathbf{r}(\hat{\mathbf{z}}) = (0.0049225,\ 0,\ -0.24612,\ 0.36919)^{\mathsf{T}}$, and the gradient (4.28), column by column of $\mathbf{J}$,

$$\mathbf{J}^{\mathsf{T}}\mathbf{r}(\hat{\mathbf{z}}) = \begin{bmatrix} -3000 \times 0.0049225 + 40 \times 0.36919 \\ 100 \times 0.0049225 + 2 \times (-0.24612) \\ 100 \times 0 \end{bmatrix} = \begin{bmatrix} -14.767 + 14.767 \\ 0.4922 - 0.4922 \\ 0 \end{bmatrix} = \mathbf{0},$$

to numerical precision: the state satisfies the first-order optimality condition. Each line reads as a balance of rows on one entry: the prior's pull on $\rho$ is held by the first physical row, which in turn is held at $T_1$ by the reading row; nothing pulls on $T_2$ but its own physical row. The information matrix (4.27) has positive eigenvalues, so the objective is strictly convex and the stationary point is its unique minimum.

**Physical consistency.** The physical rows at the fitted state, in kelvin, are $\sigma_p r_1 = 4.9 \times 10^{-5}\,\mathrm{K}$ and $\sigma_p r_2 = 0$: both halls balance to well within the tolerance, and the first imbalance shrinks with $\sigma_p$. The difference between the observed and the fitted test temperature is $23.2 - 23.0769 = 0.1231\,\mathrm{K}$, a quarter of the assumed sensor standard deviation of 0.5 K. The model is consistent with this reading. Agreement with one reading does not validate the physical model: with one reading and one free direction, the fit has nearly enough freedom to reproduce any value.

### 4.4.2 Prior predictive check

Before the reading was used, the model and the prior predicted a distribution for it. By equation (4.11), $Y = \mathbf{h}^{\mathsf{T}}\mathbf{z} + \epsilon$, with the state under its prior $\mathcal{N}(\mathbf{z}^{(0)}, \boldsymbol{\Sigma}_0)$ of Section 4.3.5 and $\epsilon$ independent of it, so $Y$ is Gaussian with

$$\mathbb{E}[Y] = \mathbf{h}^{\mathsf{T}}\mathbf{z}^{(0)} = 22.8^{\circ}\mathrm{C}, \qquad \mathrm{Var}(Y) = \mathbf{h}^{\mathsf{T}}\boldsymbol{\Sigma}_0\mathbf{h} + \sigma_T^2 = (\boldsymbol{\Sigma}_0)_{22} + \sigma_T^2 = 0.5625 + 0.25 = 0.8125\,\mathrm{K}^2,$$

the variances adding because the two terms are independent, and $(\boldsymbol{\Sigma}_0)_{22} = 30^2 \times 0.025^2$ in the limit of exact rows. Thus

$$Y \sim \mathcal{N}\big(22.8,\ 0.9014^2\big). \tag{4.70}$$

The observed reading lies $(23.2 - 22.8)/0.9014 = 0.444$ standard deviations from the predicted mean, an unremarkable distance: the reading is compatible with the prior information and the model. This check is distinct from computing the posterior. It asks whether the evidence was plausible under the assumptions before those assumptions were used to interpret it; a reading several predictive standard deviations away would have signaled a wrong prior, a wrong model or a faulty sensor before any posterior was formed.

### 4.4.3 The evidence by two routes

The evidence (4.18) of the model is the prior predictive density (4.70) evaluated at the reading,

$$p(y_T \mid M) = \mathcal{N}(y_T;\ \mathbf{h}^{\mathsf{T}}\mathbf{z}^{(0)},\ \mathbf{h}^{\mathsf{T}}\boldsymbol{\Sigma}_0\mathbf{h} + \sigma_T^2),$$

with $\mathcal{N}(y; m, v)$ the Gaussian density of Section 4.2.1. With $m = 22.8°\mathrm{C}$ and $v = 0.8125\,\mathrm{K}^2$,

$$p(y_T = 23.2 \mid M) = \frac{1}{\sqrt{2\pi \times 0.8125}} \exp\left[-\frac{(23.2 - 22.8)^2}{2 \times 0.8125}\right] = 0.44259 \times 0.90623 = 0.4011\,\mathrm{K}^{-1}. \tag{4.71}$$

The Laplace form (4.19) gives the same number from the fit. At $\sigma_p = 0.01\,\mathrm{K}$, $\det(\boldsymbol{\Sigma})^{1/2} = 1.3867 \times 10^{-6}$ and $\mathcal{L}(\hat{\mathbf{z}}) = 0.09845$ (Table 4.4), so

$$p(y_T \mid M) = 0.39894 \times \frac{\det(\boldsymbol{\Sigma})^{1/2}}{\sigma_p^2} \times \frac{\mathrm{e}^{-\mathcal{L}(\hat{\mathbf{z}})}}{\sigma_T\,\sigma_0} = 0.39894 \times 0.013867 \times 0.90624 \times 80 = 0.4011\,\mathrm{K}^{-1},$$

with $\det(\boldsymbol{\Sigma})^{1/2}/\sigma_p^2 = 0.013867$ against its limit $1/\sqrt{5200} = 0.013868$, the width of the posterior along the free direction.

**Common mistake**

Reading an evidence value as a probability that the model is correct. The density $0.4011\,\mathrm{K}^{-1}$ is not a 40 % confidence in the thermal model: it is a density, in inverse units of the reading, whose value changes by a factor of 1000 if the reading is expressed in millikelvin. A posterior probability of a model requires the evidence of the competing models and their prior probabilities. Likewise, the posterior probability 0.4267 is a confidence in the claim about $T_2$ under the model, while the evidence ratio is a confidence in the model the claim was computed under; the two answer different questions.

### 4.4.4 Posterior predictive validation with a new observation

Suppose a further, independent reading is to be taken at the same 30 kW condition. Before its value is seen, its distribution given the evidence already in hand is the posterior predictive. The new reading is $Y_{\mathrm{new}} = \mathbf{h}^{\mathsf{T}}\mathbf{z} + \epsilon_{\mathrm{new}}$, with $\mathbf{z} \mid \mathbf{y} \sim \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ and $\epsilon_{\mathrm{new}} \sim \mathcal{N}(0, \sigma_T^2)$ independent of it, so

$$Y_{\mathrm{new}} \mid \mathbf{y} \sim \mathcal{N}(\mathbf{h}^{\mathsf{T}}\hat{\mathbf{z}},\ \mathbf{h}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{h} + \sigma_T^2) = \mathcal{N}(23.0769,\ 0.17308 + 0.25) = \mathcal{N}(23.0769,\ 0.6504^2), \tag{4.72}$$

with $\mathbf{h}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{h} = \Sigma_{22}$ from equation (4.35). The predictive variance has two parts: the remaining uncertainty about the test temperature, $0.173\,\mathrm{K}^2$, and the error of the new reading itself, $0.25\,\mathrm{K}^2$. A further reading of the same sensor can therefore never be predicted more precisely than $\sigma_T$, however well the resistance becomes known.

An independent new reading compared with this distribution assesses the predictive performance of the model: a reading a fraction of a predictive standard deviation from 23.0769°C passes, and one several standard deviations away does not. Repeated large discrepancies, or a systematic sign in them, would indicate that the model, the prior, the measurement-error assumptions or the representation of the operating condition need revision.

### 4.4.5 The limit of exact rows

Table 4.9 repeats the fit at three tolerances and in the limit $\sigma_p \to 0$. Each column is one tolerance; the rows give the MAP resistance and full-load temperature, their standard deviations, the correlation of the two hall temperatures, the threshold probability of Section 4.3.5, and the

condition number of the information matrix. By $\sigma_p = 0.01\,\mathrm{K}$ the full-load temperature and the threshold probability have stopped moving at the printed precision; the resistance and its spread still move in their sixth decimal, the full-load spread in its fourth; at $\sigma_p = 0.001\,\mathrm{K}$ all have stopped. The fit at $\sigma_p = 0.001\,\mathrm{K}$ reproduces the limit values of Section 4.2.10 to every printed digit.

Table 4.9: The fit of the hall state as the model tolerance of the physical rows shrinks. The last column is the limit of exact rows.

| Quantity | $\sigma_p = 0.1\,\mathrm{K}$ | $0.01\,\mathrm{K}$ | $0.001\,\mathrm{K}$ | $\sigma_p \to 0$ |
|---|---|---|---|---|
| $\hat{\rho}$ [K/kW] | 0.169119 | 0.169230 | 0.169231 | 0.169231 |
| $\sqrt{\Sigma_{11}}$ [K/kW] | 0.014056 | 0.013869 | 0.013868 | 0.013868 |
| $\hat{T}_2$ [°C] | 28.1471 | 28.1538 | 28.1538 | 28.1538 |
| $\sqrt{\Sigma_{33}}$ [K] | 0.84926 | 0.83223 | 0.83205 | 0.83205 |
| correlation of $T_1$, $T_2$ [–] | 0.96522 | 0.99964 | 0.999996 | 1 |
| $\Pr(T_2 < 28°\mathrm{C})$ [–] | 0.4312 | 0.4267 | 0.4267 | 0.4267 |
| condition number of $\mathbf{J}^\mathsf{T}\mathbf{J}$ [–] | $4.0 \times 10^5$ | $3.9 \times 10^7$ | $3.9 \times 10^9$ | – |

The $\sqrt{\Sigma_{11}}$ row widens with the tolerance as equation (4.33) requires, the tolerance acting on the reading as extra sensor variance, and the limit column is the matrix of equation (4.35); its entry for $\sqrt{\Sigma_{11}}$ is the square root of $\mathbf{\Sigma}_{\mathrm{red}}$, the covariance of the reduced state (Section 4.2.11). The fitted covariances satisfy equation (4.34) at every tolerance of the table ($0.69260 = 0.69250 + 0.0001$ at $\sigma_p = 0.01\,\mathrm{K}$, $0.72125 = 0.71125 + 0.01$ at $0.1\,\mathrm{K}$). A tighter tolerance than $0.001\,\mathrm{K}$ buys nothing and costs conditioning: the condition number grows a hundredfold per decade, and at $\sigma_p = 10^{-6}\,\mathrm{K}$ it reaches $4 \times 10^{15}$ and round-off already moves the fourth digit of $\sqrt{\Sigma_{33}}$.

### 4.4.6 Second routes to the full-load answers

**The propagated full-load temperature.** As the third entry, with $\mathbf{g} = (0,\ 0,\ 1)^\mathsf{T}$, the full-load temperature has $\mathbf{g}^\mathsf{T}\mathbf{\Sigma}\mathbf{g} = \Sigma_{33} = 0.69231\,\mathrm{K}^2$; as the prediction $18{+}60\rho$, with $\mathbf{g} = (60,\ 0,\ 0)^\mathsf{T}$, it has $\mathbf{g}^\mathsf{T}\mathbf{\Sigma}\mathbf{g} = 60^2\,\Sigma_{11} = 3600/5200 = 0.69231\,\mathrm{K}^2$, the same number, as Section 4.3.1 requires in the limit.

**The threshold through the resistance.** By equation (4.58), $t^* = (28 - 28.1538)/60 = -0.0025641\,\mathrm{K/kW}$ and

$$\rho^* = \hat{z}_1 + t^* = 0.169231 - 0.002564 = 0.166667\,\mathrm{K/kW},$$
$$\Pr(\rho < \rho^* \mid \mathbf{y}) = \Phi\!\left(\frac{-0.002564}{0.013868}\right) = \Phi(-0.1849) = 0.4267,$$

the number of equation (4.61).

### 4.4.7 The propagation check against draws

For a linear output the delta method is exact, and the check is the algebraic agreement of the entry and the gradient already shown in Section 4.4.6. For a nonlinear output the linearized result must be compared with the full nonlinear evaluation at draws from the posterior. The procedure is the following.

1. Factor the covariance as $\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^\mathsf{T}$ with $\mathbf{L}$ lower triangular (the Cholesky factor, which exists at any finite tolerance), and draw $N$ independent states $\mathbf{z}^{(j)} = \hat{\mathbf{z}} + \mathbf{L}\boldsymbol{\varepsilon}^{(j)}$, $j = 1, \ldots, N$,

each $\boldsymbol{\varepsilon}^{(j)}$ a vector of three independent standard normal numbers, so that $\mathbf{z}^{(j)} \sim \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$. At $\sigma_p = 0.01\,\mathrm{K}$,

$$\mathbf{L} = \begin{bmatrix} 0.013869 & 0 & 0 \\ 0.41592 & 0.0099980 & 0 \\ 0.83217 & 0 & 0.010000 \end{bmatrix},$$

whose first column is $\sqrt{\Sigma_{11}}$ times (nearly) $\mathbf{v}$ and whose other two columns have the width $\sigma_p$ of the physical rows. (In the limit of exact rows the factor has the single column $\mathbf{v}/\sqrt{5200}$, and a draw is $\hat{\mathbf{z}} + \mathbf{v}\,\varepsilon^{(j)}/\sqrt{5200}$, along the free direction.)

2. Evaluate the output at each drawn state through the full model, for example $T_2^{(j)} = z_3^{(j)}$ or $G^{(j)} = 1/z_1^{(j)}$.
3. Estimate a threshold probability as the fraction of draws satisfying the event, $\hat{P} = N^{-1}\sum_j \mathbf{1}[f(\mathbf{z}^{(j)}) < f^*]$, whose standard error is $\sqrt{\hat{P}(1-\hat{P})/N}$; estimate quantiles as the empirical quantiles of the $f(\mathbf{z}^{(j)})$.
4. Compare with the closed forms of the delta method. Agreement within the standard error means the closed form can be trusted.

A few hundred draws resolve a probability near one half to about $\pm 0.03$; $N = 10^5$ resolves it to about $\pm 0.002$.

Table 4.10 applies the procedure to the full-load temperature, which is linear in $\rho$, and to the conductance $G = 1/\rho$, which is not. The first column names the quantity; the second gives the delta-method value, from a Gaussian with mean $f(\hat{\mathbf{z}})$ and standard deviation $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$; the third gives the exact value of the push-forward of the posterior, the limit to which the draw estimates converge, computed here in closed form because both outputs are monotone in $\rho$ (the $q$-quantile of $G$ is $1/(\hat{\rho} - z_q\sqrt{\Sigma_{11}})$ with $z_q$ the standard normal quantile); the fourth gives one run of $N = 10^5$ drawn states at $\sigma_p = 0.01\,\mathrm{K}$, generated with a seeded pseudo-random generator so that the run can be repeated exactly. The conductance carries the gradient of Section 4.3.6.

Table 4.10: The propagation check: delta method against the exact push-forward of the posterior, whose first entry is $\mathcal{N}(0.169231, 0.013868^2)$, and one run of $10^5$ states drawn as $\hat{\mathbf{z}} + \mathbf{L}\boldsymbol{\varepsilon}$ (standard error of a probability about 0.0015). The threshold $G = 6.25\,\mathrm{kW/K}$ is the handbook conductance.

| Quantity | Delta method | Exact | $10^5$ draws |
|---|---|---|---|
| $T_2$, median [°C] | 28.154 | 28.154 | 28.152 |
| $T_2$, 5 % and 95 % [°C] | 26.785, 29.522 | 26.785, 29.522 | 26.786, 29.516 |
| $\Pr(T_2 < 28°\mathrm{C})$ [–] | 0.4267 | 0.4267 | 0.4269 |
| $G$, median [kW/K] | 5.909 | 5.909 | 5.910 |
| $G$, 5 % and 95 % [kW/K] | 5.113, 6.706 | 5.207, 6.830 | 5.210, 6.828 |
| $\Pr(G > 6.25\,\mathrm{kW/K})$ [–] | 0.2407 | 0.2528 | 0.2527 |

For the temperature the three columns agree to within the sampling error, and the closed form needs no correction. For the conductance the median is right in all three columns, since $1/\rho$ is monotone and maps the median of $\rho$ to the median of $G$, but the delta-method band is shifted: the exact band $[5.207, 6.830]$ is skewed to the right of $\hat{G}$, while the linearized band $[5.113, 6.706]$ is symmetric about it, and the delta method understates the probability that the conductance exceeds its handbook value by about 0.012. The draws reproduce the exact column to within their standard error. Figure 4.7 shows the cause. The horizontal axis is the conductance and the vertical axis its density; the solid curve is the exact density of $G = 1/\rho$, obtained by the change of variables $p_G(G) = p(\rho = 1/G \mid \mathbf{y})/G^2$ of Section 4.3.3, and the dashed curve the delta-method Gaussian. The exact density has its mode to the left of 5.909 and a longer right tail, and the Gaussian cannot represent the asymmetry.

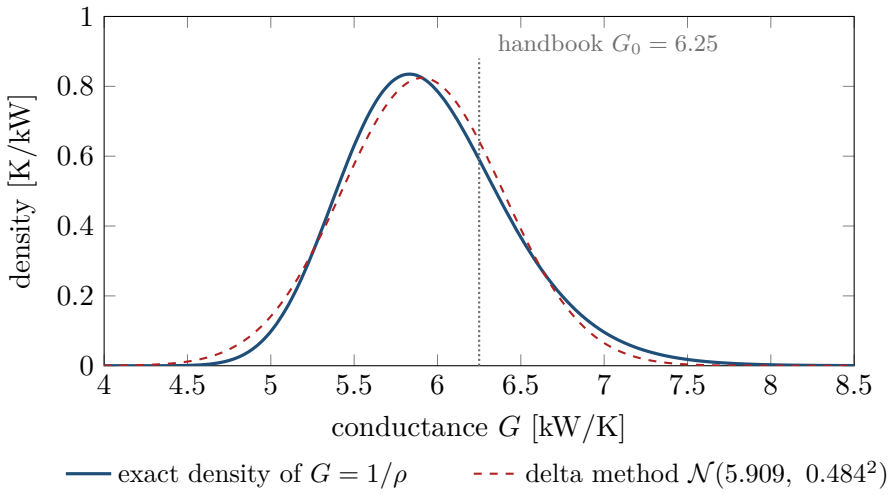



Figure 4.7: Posterior density of the conductance $G = 1/\rho$ under $\rho \mid \mathbf{y} \sim \mathcal{N}(0.169231, 0.013868^2)$: exact, by change of variables, $p_G(G) = \varphi((1/G - 0.169231)/0.013868)/(0.013868\, G^2)$ with $\varphi$ the standard normal density; and the delta-method Gaussian $\mathcal{N}(1/\hat{\rho}, \Sigma_{11}/\hat{\rho}^4)$. Dotted line: the handbook conductance.

The conductance also shows why "the posterior mean of $G$" must be handled with care. Under a Gaussian posterior for $\rho$, the mean of $1/\rho$ does not exist in the strict sense, because the Gaussian has positive density at $\rho = 0$, where $1/\rho$ is not integrable; the mass involved here is about $10^{-34}$ and has no practical effect, but it is a reminder that moments of a transformed quantity are properties of the parameterization. Quantiles and threshold probabilities, which the push-forward preserves exactly for a monotone output, are the robust summaries, and they are what the check compares.

### 4.4.8 Sensitivity to the value read

Table 4.11 repeats the assessment for other values of the single test reading. The first numeric column is the prior; each further column is the posterior after a different value of the reading, the actual one, 23.2°C, among them. The rows give the MAP resistance from equation (4.29), the mean and standard deviation of the full-load temperature, and the probability (4.57). The standard deviation is the same for every reading, and the mean moves along the gain vector (4.59), which with the limit covariance (4.35) is

$$\frac{\partial \hat{\mathbf{z}}}{\partial y_T} = \frac{\boldsymbol{\Sigma}\mathbf{h}}{\sigma_T^2} = 4 \begin{bmatrix} 0.0057692 \\ 0.17308 \\ 0.34615 \end{bmatrix} = \begin{bmatrix} 0.023077 \\ 0.69231 \\ 1.38462 \end{bmatrix},$$

four times the second column of the limit covariance, in K/kW per K and K per K. One column of the table, for $y_T = 22.4$°C, is the actual state moved by $-0.8$ K along it: $\hat{\rho} = 0.169231 - 0.8 \times 0.023077 = 0.150769$, $\mathbb{E}[T_2 \mid \mathbf{y}] = 28.1538 - 0.8 \times 1.38462 = 27.046$°C, and $\Phi((28 - 27.046)/0.8321) = \Phi(1.1464) = 0.8742$.

A reading of 22.8°C, equal to what the prior predicted, leaves the mean where it was and only narrows the distribution; because the mean then lies below the limit, narrowing alone raises the probability from 0.605 to 0.685. The observed 23.2°C moves the mean across the limit and lowers the probability to 0.427. The posterior probability equals the prior one for a reading of 22.93°C and falls below one half for any reading above 23.09°C.

Table 4.11: The full-load assessment before the test and after a single test reading of various values ($\sigma_T = 0.5\,\mathrm{K}$, prior $\mathcal{N}(0.160, 0.025^2)$). The column 23.2°C is the observed reading.

| Quantity | Prior | $y_T = 22.4$ | 22.8 | 23.2 | 23.6 | 24.0 |
|---|---|---|---|---|---|---|
| $\hat{\rho}$ [K/kW] | 0.16000 | 0.15077 | 0.16000 | 0.16923 | 0.17846 | 0.18769 |
| $\mathbb{E}[T_2]$ [°C] | 27.60 | 27.05 | 27.60 | 28.15 | 28.71 | 29.26 |
| $\mathrm{sd}(T_2)$ [K] | 1.500 | 0.832 | 0.832 | 0.832 | 0.832 | 0.832 |
| $\Pr(T_2 < 28°\mathrm{C})$ [–] | 0.6051 | 0.8742 | 0.6846 | 0.4267 | 0.1975 | 0.0647 |

## 4.5 The report (step 6)

One reading at the test condition overturns the handbook's answer. The handbook resistance put the hall at 27.6°C at full load, under the limit, with probability 0.6051 of staying below it; the reading of 23.2°C moves the resistance up, from 0.160 to 0.169 K/kW, and leaves the probability at 0.4267, near even odds, although it narrows the full-load spread from 1.5 to 0.832 K. The answers to the questions of Table 4.2 are given in the units in which the questions were asked, conditional on the steady-state model, the test reading and the handbook prior; the release decision follows from them.

### 4.5.1 The answers and their bands

Table 4.12 gives the answers. Each row is a question; the second column gives the answer in the limit of exact physical rows, except the last, at the tolerance of Section 4.3.8, and the third its band: the posterior standard deviation of an estimate, with the central 95 % credible interval of the full-load temperature, for the probability the range over the model tolerances of Table 4.9, and for the set point the back-off it already contains.

Table 4.12: The report for the equipment hall at 60 kW, after the test reading of 23.2°C. Answers in the limit of exact physical rows, the set point at $\sigma_p = 0.01\,\mathrm{K}$; bands are posterior standard deviations, a credible interval, the range over model tolerances $\sigma_p$ from 0 to 0.1 K, or the back-off.

| Control engineer's question | Answer | Band |
|---|---|---|
| What is the thermal resistance of the air path? | 0.16923 [K/kW]; conductance 5.909 [kW/K] | standard deviation 0.01387 [K/kW]; conductance 5–95 % band 5.113 to 6.706 [kW/K], exact 5.207 to 6.830 [kW/K] |
| What was the hall temperature during the test? | 23.077 [°C] | standard deviation 0.416 [K] |
| What will the hall temperature be at full load? | 28.15 [°C] | standard deviation 0.83 [K]; 26.52 to 29.78 [°C], central 95 % interval |
| Can the hall be released for 60 kW within 28°C? | probability 0.427 [–]; of exceeding, 0.573 [–] | 0.4267 to 0.4312 [–] over the model tolerances |
| How far must the set point move to hold it at 0.95? | from 18 to $r_{95} = 16.4773$ [°C] | 16.4560 to 16.4776 [°C] over the model tolerances; moves 1.5227 [K]: 0.1538 for the mean, 1.3689 for the spread |

The observed temperature was slightly warmer than the prior model predicted (22.8°C, the second entry of $\mathbf{z}^{(0)}$), so the inferred resistance rose above its handbook value and the implied conductance fell below 6.25 kW/K. The inferred test temperature is not the observed 23.2°C: the reading is noisy, and the prior, carried to $T_1$ by the physical row, also contributes

information. The conductance in the first row is $1/\hat{\rho}$, the conductance evaluated at the estimated resistance; it is not the posterior mean of the conductance (Section 4.4.7). With correlations above 0.9996 the three entries are known almost only as one combination, the free direction $\mathbf{v}$, and the full-load interval contains temperatures both below and above the limit.

### 4.5.2 The release decision and the entry in the commissioning log

Conditional on the steady-state model, the reading and the prior, the estimated full-load temperature lies slightly above the limit, with substantial uncertainty on both sides of it. The reading reduced the uncertainty and, at the same time, lowered the probability of staying within the acceptance limit, from 0.605 before it to 0.427. The evidence does not support a high-confidence claim that the requested load keeps the hall within the limit: the posterior probability of doing so is about 0.43. A decision to raise the load to full would require an assurance level $\kappa$, stated explicitly, and a consideration of the consequences of exceeding the limit. Further readings or more cooling capacity may be appropriate, but their value should be computed, not assumed. That is what an inference report gives the commissioning record: the physical interpretation of a mathematical result, its uncertainty, and its implications for a decision.

Table 4.11 shows that the direction of the change was set by the value read, not by the act of reading. New evidence does not necessarily make the plant look safer: it can make it look less safe, or move the probability in either direction, depending on the value observed. The purpose of inference is to improve the knowledge of the plant, not to produce reassuring probabilities.

> **Principle 4.2.** *A reading narrows the posterior by an amount fixed before it is taken, but the direction in which it moves a threshold probability is set by the value it returns. Evidence improves knowledge; it does not promise reassurance.*

The entry in the commissioning log states the level it uses. At an assurance level of 0.95 this record does not support release for 60 kW at the 18°C set point: the full-load temperature is about 28.15°C, between 26.52 and 29.78°C, and the probability of staying below 28°C is 0.4267. The level would be met at a set point of 16.4773°C, which contains a back-off of 1.3689 K, provided the discharge controller reaches that value at full load and the resistance holds there, neither of which the test exercised. The verdict rests on one record, one draw of the sensor's error, and Table 4.11 shows how far another value would have moved it.

## 4.6 The next reading (step 7)

The report leaves the release question near even odds, with a band that a single reading at the test condition has set. A further reading is one more row of the residual vector, and the information it adds is known before its value is.

### 4.6.1 The preposterior covariance

A further reading at the test condition is the row $(T_1 - y)/\sigma_T$, whose line of the Jacobian is $\mathbf{j} = (0,\ 2,\ 0)$, the line of the first reading; a reading at full load, by a sensor of the same spread, is the row $(T_2 - y)/\sigma_T$, with line $\mathbf{j}_{\mathrm{f}} = (0,\ 0,\ 2)$. The information that a reading adds, $\mathbf{j}^{\mathsf{T}}\mathbf{j}$, is known before the reading is taken, because its Jacobian depends on the sensor and on where it is placed, not on the value. For a candidate reading $a$ with information $\mathbf{\Lambda}_a$ the preposterior covariance is

$$\mathbf{\Sigma}_a = \left(\mathbf{\Sigma}^{-1} + \mathbf{\Lambda}_a\right)^{-1}, \tag{4.73}$$

formed at a finite tolerance, where $\mathbf{\Sigma}^{-1} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ exists, the object on which candidate readings are ranked. In the limit of exact rows the posterior stays on the free direction, and a candidate adds

to the information $\|\mathbf{J}\mathbf{v}\|^2 = 5200$ along it the square of its own line times $\mathbf{v}$: $(\mathbf{j}\mathbf{v})^2 = 60^2 = 3600$ for a reading at 30 kW and $(\mathbf{j}_\mathrm{f}\mathbf{v})^2 = 120^2 = 14\,400$ for a reading at 60 kW. The full-load standard deviation after the candidate is 60 divided by the square root of the total. For $n$ further readings at 30 kW, $\mathbf{\Lambda}_a = n\,\mathbf{j}^\mathsf{T}\mathbf{j}$, which adds $4n$ to the $(T_1, T_1)$ entry, and the third diagonal entry of $\mathbf{\Sigma}_a$ gives the standard deviation of the full-load temperature: 0.640 K for $n = 1$, 0.539 K for $n = 2$, 0.429 K for $n = 4$ and 0.325 K for $n = 8$. More readings resolve the question; they do not decide its answer. If the full-load temperature truly lies above 28°C, the probability tends to zero as the readings accumulate, and the decision then rests on the air path, not on the readings.

### 4.6.2 The candidate readings

Table 4.13 ranks three candidates. Each row is a reading the engineer can request; the columns give the information along $\mathbf{v}$ after it, the standard deviations of the resistance and of the full-load temperature that the preposterior covariance (4.73) predicts, and the width it removes from the current full-load standard deviation of 0.8321 K. One reading at full load carries four times the information of one at the test condition, because the full-load temperature moves twice as far as the test temperature along $\mathbf{v}$, and it removes exactly the width of four test readings.

Table 4.13: Candidate readings for the hall, in the limit of exact physical rows, from the posterior after the test reading. Each reading has $\sigma_T = 0.5$ K; the width removed is taken from the current $\sqrt{\Sigma_{33}} = 0.8321$ K.

| Candidate | Information along $\mathbf{v}$ [(kW/K)$^2$] | sd($\rho$) after [K/kW] | sd($T_2$) after [K] | Width removed [K] |
|---|---|---|---|---|
| One further reading at 30 kW | 8800 | 0.01066 | 0.640 | 0.192 |
| Four further readings at 30 kW | 19 600 | 0.00714 | 0.429 | 0.403 |
| One reading at 60 kW | 19 600 | 0.00714 | 0.429 | 0.403 |

### 4.6.3 The recommendation

Four further readings at the test condition and one reading at full load narrow the answer equally, and the choice between them rests on what else each tests. Readings at 30 kW need no operation at the requested load, but repeated readings of one sensor share its bias, and no number of them can show whether the resistance at 60 kW is the resistance at 30 kW. A reading at full load tests exactly that assumption, on which the prediction rests (Section 4.6.6), at the cost of running the hall at the load in question. The next reading is therefore one reading at full load, taken in a supervised run, the load raised in steps and held at 60 kW until the hall settles, which the engineer can stop if the hall approaches its limit; where such a run is not acceptable, it is four further readings at the test condition, by more than one sensor where possible. A permanent second hall sensor is Chapter 1's recommendation; here it is one more row of this kind. Chapter 16 ranks the readings available on the two-unit hall by their preposterior in the same way.

### 4.6.4 A second observation as an update

Before any run at full load, the first candidate of Table 4.13 arrives: a second, independent reading at the test condition, $y_{T,2} = 23.0$°C, with the same standard deviation $\sigma_T = 0.5$ K; the first reading is now written $y_{T,1} = 23.2$°C. Against the posterior predictive (4.72) it lies $(23.0 - 23.0769)/0.6504 = -0.12$ standard deviations from the predicted mean and passes that check. The posterior after the first reading serves as the prior for the second.

**The precision form.** Information from independent rows adds. The new reading is one more row of the residual vector, $(T_1 - y_{T,2})/\sigma_T$, whose line of the Jacobian is $\mathbf{j} = (0,\ 2,\ 0)$, the same as the first reading's. Its contribution to the information matrix is the outer product $\mathbf{j}^\mathsf{T}\mathbf{j}$, which changes the $(T_1, T_1)$ entry alone,

$$\mathbf{\Lambda}_2 = \mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{j}^\mathsf{T}\mathbf{j} = \mathbf{J}^\mathsf{T}\mathbf{J} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{\Sigma}_2 = \mathbf{\Lambda}_2^{-1},$$

the matrix (4.68). The old rows are balanced at $\hat{\mathbf{z}}$, $\mathbf{J}^\mathsf{T}\mathbf{r}(\hat{\mathbf{z}}) = \mathbf{0}$, so the gradient of the enlarged objective at $\hat{\mathbf{z}}$ comes from the new row alone, $\mathbf{j}^\mathsf{T}(23.0769 - 23.0)/0.5 = (0,\ 0.30769,\ 0)^\mathsf{T}$, and one Gauss–Newton step (4.29) from $\hat{\mathbf{z}}$ gives $\hat{\mathbf{z}}_2 = \hat{\mathbf{z}} - 0.30769\,(\mathbf{\Sigma}_2)_{\cdot 2}$, with $(\mathbf{\Sigma}_2)_{\cdot 2}$ the second column of $\mathbf{\Sigma}_2$. In the limit of exact rows this column is $(3.4091 \times 10^{-3},\ 0.10227,\ 0.20455)^\mathsf{T}$, and

$$\hat{\mathbf{z}}_2 = \begin{bmatrix} 0.169231 \\ 23.0769 \\ 28.1538 \end{bmatrix} - 0.30769 \begin{bmatrix} 0.0034091 \\ 0.10227 \\ 0.20455 \end{bmatrix} = \begin{bmatrix} 0.168182 \\ 23.0455 \\ 28.0909 \end{bmatrix}.$$

The first entry of $\mathbf{\Sigma}_2$ is $1.1364 \times 10^{-4}\,(\mathrm{K/kW})^2$, so that $\sqrt{(\mathbf{\Sigma}_2)_{11}} = 0.010660\,\mathrm{K/kW}$.

**The rank-one form.** The same result follows from the first posterior alone, without forming and inverting $\mathbf{\Lambda}_2$. The second reading measures $\mathbf{h}^\mathsf{T}\mathbf{z} = T_1$ with $\mathbf{h} = (0,\ 1,\ 0)^\mathsf{T}$; its predicted value under the first posterior is $\mathbf{h}^\mathsf{T}\hat{\mathbf{z}} = 23.0769°\mathrm{C}$ and its predictive variance, from equation (4.72), is $S = \mathbf{h}^\mathsf{T}\mathbf{\Sigma}\mathbf{h} + \sigma_T^2 = \Sigma_{22} + \sigma_T^2 = 0.42308\,\mathrm{K}^2$. Conditioning the Gaussian on the new row gives the update

$$\hat{\mathbf{z}}_2 = \hat{\mathbf{z}} + \mathbf{K}\,(y_{T,2} - \mathbf{h}^\mathsf{T}\hat{\mathbf{z}}), \qquad \mathbf{\Sigma}_2 = \mathbf{\Sigma} - S\,\mathbf{K}\mathbf{K}^\mathsf{T}, \qquad \mathbf{K} = \frac{\mathbf{\Sigma}\mathbf{h}}{S}, \tag{4.74}$$

where $\mathbf{K}$ is the gain vector, one entry per state entry, and $y_{T,2} - \mathbf{h}^\mathsf{T}\hat{\mathbf{z}}$ the innovation, the part of the reading the first posterior did not predict. Because $\mathbf{\Sigma}_2\mathbf{h} = \mathbf{\Sigma}\mathbf{h} - \mathbf{K}\,\mathbf{h}^\mathsf{T}\mathbf{\Sigma}\mathbf{h} = (\sigma_T^2/S)\,\mathbf{\Sigma}\mathbf{h}$, the gain is also $\mathbf{K} = \mathbf{\Sigma}_2\mathbf{h}/\sigma_T^2$, the same vector as the gain (4.59) of the posterior that includes the new reading. The gain follows from writing the posterior mean as $\hat{\mathbf{z}} + \mathrm{Cov}(\mathbf{z}, Y_{\mathrm{new}})\,(y_{T,2} - \mathbb{E}[Y_{\mathrm{new}}])/S$ for the jointly Gaussian state and new reading, with $\mathrm{Cov}(\mathbf{z}, Y_{\mathrm{new}}) = \mathbf{\Sigma}\mathbf{h}$, the second column of $\mathbf{\Sigma}$. With the limit covariance (4.35),

$$\mathbf{K} = \frac{1}{0.42308} \begin{bmatrix} 0.0057692 \\ 0.17308 \\ 0.34615 \end{bmatrix} = \begin{bmatrix} 0.013636 \\ 0.40909 \\ 0.81818 \end{bmatrix},$$

in $\mathrm{kW}^{-1}$ for the first entry and dimensionless for the other two; the innovation is $23.0 - 23.0769 = -0.0769\,\mathrm{K}$, and $\hat{\mathbf{z}}_2 = \hat{\mathbf{z}} - 0.0769\,\mathbf{K} = (0.168182,\ 23.0455,\ 28.0909)^\mathsf{T}$, the state of the precision form. The third entry of $\mathbf{K}$ is the limit of the gain 0.81806 of Section 4.3.8, and at this reading of 23.0°C the set point at 0.95 is the 16.8566°C of Table 4.8. The covariance update subtracts the outer product $S\mathbf{K}\mathbf{K}^\mathsf{T}$; on its first and last diagonal entries,

$$(\mathbf{\Sigma}_2)_{11} = 1.9231 \times 10^{-4} - 0.42308 \times 0.013636^2 = 1.1364 \times 10^{-4}\,(\mathrm{K/kW})^2,$$
$$(\mathbf{\Sigma}_2)_{33} = 0.69231 - 0.42308 \times 0.81818^2 = 0.40909\,\mathrm{K}^2,$$

the values of the precision form. Because $S\,\mathbf{K}^\mathsf{T} = \mathbf{h}^\mathsf{T}\mathbf{\Sigma}$, the covariance update is also

$$\mathbf{\Sigma}_2 = (\mathbf{I} - \mathbf{K}\mathbf{h}^\mathsf{T})\,\mathbf{\Sigma},$$

with $\mathbf{I}$ the identity matrix: each column of $\mathbf{\Sigma}$ loses the gain vector times that column's second entry. The correction is the outer product of one vector, which is the reason for the name rank-one update. For a linear-Gaussian model it is exact, and one more independent reading costs one product of $\mathbf{\Sigma}$ with a vector rather than a new inference. For a strongly nonlinear model, or after a shift of the estimate large enough to make the Jacobian at the old estimate stale, the model is linearized again or the full fit is repeated.

**The updated assessment.** Table 4.14 follows the assessment through the two readings. Each column is a state of knowledge, before any reading, after the first, after both; the rows are the information about the resistance, the reciprocal of the first diagonal entry of the covariance, the MAP resistance, the mean and standard deviation of the full-load temperature, the threshold probability, and the 95 % credible interval.

Table 4.14: The full-load assessment before the test, after the first reading (23.2°C) and after the second (23.0°C).

| Quantity | Prior | One reading | Two readings |
|---|---|---|---|
| $1/\Sigma_{11}$ [(kW/K)$^2$] | 1600 | 5200 | 8800 |
| $\hat{\rho}$ [K/kW] | 0.160000 | 0.169231 | 0.168182 |
| $\mathbb{E}[T_2]$ [°C] | 27.600 | 28.154 | 28.091 |
| sd($T_2$) [K] | 1.500 | 0.832 | 0.640 |
| Pr($T_2 < 28$°C) [–] | 0.6051 | 0.4267 | 0.4435 |
| 95 % interval [°C] | 24.66, 30.54 | 26.52, 29.78 | 26.84, 29.34 |

After the second reading the full-load temperature is the third entry of $\hat{\mathbf{z}}_2$, 28.091°C, with standard deviation $\sqrt{(\boldsymbol{\Sigma}_2)_{33}} = \sqrt{0.40909} = 0.640$ K, and the probability of staying below the limit becomes $\Phi(-0.1421) = 0.4435$. The interval narrowed by a quarter, yet the probability moved only from 0.427 to 0.443, because two effects acted in opposite directions. The narrowing alone, with the mean held at 28.154°C, would have lowered the probability to 0.405, since the mean lies above the limit; the shift of the mean alone, with the spread held at 0.832 K, would have raised it to 0.456. The effect of actual new evidence cannot be described by its reduction of the covariance alone: the value observed matters.

### 4.6.5 What the analysis showed

The handbook said the hall passes at 27.6°C, and the reading scaled to full load said it fails at 28.4°C; the analysis says neither. With the reading, the handbook and the physics combined, the probability that the hall stays below 28°C at 60 kW is 0.4267, less than even, and the full-load temperature is not known well enough to call it either way. One reading at the test condition fixes the thermal resistance of the air path to about ±8 % and the full-load temperature to about ±0.8 K, one standard deviation each, although the hall has never run at 60 kW: the physics carries the reading along the free direction $\mathbf{v}$, from $T_1$ to $\rho$ and from $\rho$ to $T_2$. The posterior is a ridge along that direction, the three entries known almost only as one combination, and the reduced state is that ridge named by its one free entry: every question asked of the hall is a question about where it sits on the ridge. The 60 kW question in particular is a resistance question: the hall stays below 28°C exactly when the resistance stays below $\rho^*$, and the reading at 30 kW places the resistance close to $\rho^*$, which leaves the question near even odds. Where the sibling volume's analysis ends on that probability, the loop gives this one a further answer: the set point that holds the hall at 0.95 is 16.4773°C, and 1.3689 K of the way down to it is the price of the spread, not of the mean. Further readings at the test condition narrow the ridge and resolve the question without deciding its answer. A reading at full load would act on the third column of $\mathbf{J}$ directly, and its worth is what its preposterior covariance $\boldsymbol{\Sigma}_a$ says before it is taken: as much as four readings at the test condition, with the shared resistance itself put to the test.

### 4.6.6 Limits of the model

The steady-state balance (4.2) neglects the heat stored in the hall and its equipment, and it describes the temperature long after a load change, not minutes after it; Exercise 4.6 formulates the transient model that a question about the first minutes of full load would need. The single effective path assumes that the resistance is the same at 60 kW as at 30 kW; a path whose resistance changes with load, for example because the unit's fans or its coil change state, would break the link drawn in Figure 4.2, and one test reading cannot detect it. The supply air is pinned at 18°C by the discharge controller, which holds it only while the unit has capacity to spare; a unit at its limit at full load lets the supply air rise, which shifts the full-load temperature up, not only widens it; an uncertain supply air adds its variance through the budget (4.54). The set point of Section 4.3.8 leans on the same pin harder: the test never ran at 16.4773°C, the discharge controller is taken to reach it at 60 kW, and the resistance is taken not to change on the way there. The back-off there is the spread of this model and no more; the loop itself enters in Chapter 5.

The probabilistic assumptions have limits of their own. The Gaussian prior on a positive resistance is harmless here, with a negative mass of $8 \times 10^{-11}$, but a wide prior on a small quantity would place real probability below zero; a logarithmic parameterization avoids this at the price of a nonlinear model and an approximate posterior. The sensor error is taken as Gaussian, unbiased and of known spread; a bias would shift every inferred quantity, and Section 4.2.2 names the comparison of models that would detect it. The physical rows are taken in the limit of zero tolerance; a real model discrepancy is a finite $\sigma_p$, which widens $\Sigma_{11}$ and adds $\sigma_p^2$ to the full-load variance by equation (4.34), and the reduced state of Section 4.2.8 is exact only in that limit. The record is constructed: the reading of 23.2°C is a stated example value, not field data, and the chapter has no truth to check its answers against; from Chapter 5 on the readings are constructed from a stated truth and the checks compare the answers with it. Each of these is a statement about the model, and the probability 0.4267 is conditional on all of them.

Chapter 5 closes the loop that this chapter took as ideal: the hall gains a thermal capacity, a coil with a valve and a sampled controller, and the set point becomes a pin whose actuator is free; the method is unchanged.

## 4.7 What to retain

**Definition 4.1** (The anatomy of an inference)**.** A probability over a plant graph is obtained in two operations. Inference sums one term per row into the objective $\mathcal{L} = -\ln p(\mathbf{z} \mid \mathbf{y})$ up to a constant, and uses only its gradient and Hessian. A Gaussian row's term is half its squared residual; a count row's is its negative log-probability, which joins the Jacobian through its working residual and line, reproducing the row's gradient and information but not its value. The physical rows, the readings and the priors, each divided by its standard deviation, stack into one residual vector $\mathbf{r}(\mathbf{z})$ with Jacobian $\mathbf{J}$; the normal equations $(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$ lead to the MAP state $\hat{\mathbf{z}}$, and the inverse of the information matrix, $\boldsymbol{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$, is the posterior covariance, exactly for a linear-Gaussian model and under the Laplace approximation, in its Gauss–Newton form, otherwise. An exact physical relation is the limit of a row's tolerance, never a substitution; in that limit the state reduces to its free entries, with Jacobian $\mathbf{J}_{\text{rows}}\mathbf{N}$ on the null space $\mathbf{N}$ of the physical rows. Query evaluation then reads the answer off the posterior: an entry of the state with its variance $\Sigma_{ii}$, or a derived quantity $f(\mathbf{z})$ with variance $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$, and a threshold probability as the normal tail $\Phi\big((f^* - f(\hat{\mathbf{z}}))/\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}\big)$, which for an entry is $\Phi\big((f^* - \hat{z}_i)/\sqrt{\Sigma_{ii}}\big)$. A scenario node pins an input at a value not yet applied; a pinned input that enters additively moves the mean and leaves the spread alone, one that multiplies an uncertain entry, as $Q$ does, scales it, and for the set point the threshold inverts to the largest

set point, the limit less the mean rise and the back-off.

Three conclusions carry beyond the example. Physical relations connect the quantities that are measured to those that are not, which is what allows an unobserved quantity, here a full-load temperature never reached, to be inferred from indirect evidence. Uncertainty is not an error bar attached to a point estimate but a probability distribution, and it must be propagated through the physical model to answer an engineer's question. And the posterior is reusable: once the joint posterior is in hand, many questions are answered without repeating the inference, and new evidence updates it rather than replacing it. A derived quantity is propagated in one of two ways: its spread by the gradient, $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$, and its density by the change of variables, which multiplies the density by the change-of-variable factor of the map and so keeps the skew and the bounds that a linearization loses.

The inference of the hall returns the posterior state $\hat{\mathbf{z}}$ of equation (4.49), its covariance $\boldsymbol{\Sigma}$ of equations (4.50) and (4.35), and the Jacobian $\mathbf{J}$ of equation (4.25), from which the gradients of derived quantities are formed. Table 4.15 lists what the chapter has computed from them without repeating the inference: the five answers of Table 4.2, and three further quantities that later chapters reuse. Each row is a quantity, the calculation on the posterior that gives it, and the place where it is carried out.

Table 4.15: The five answers of Table 4.2 and three further quantities reused later, all computed from one posterior of the equipment hall.

| Quantity | Calculation on the posterior | Section |
|---|---|---|
| *The answers* | | |
| Thermal resistance of the air path | the entry $\hat{z}_1$ and $\sqrt{\Sigma_{11}}$ | 4.3.6 |
| Hall temperature at the test condition | the entry $\hat{z}_2$ and $\sqrt{\Sigma_{22}}$ | 4.3.6 |
| Hall temperature at full load | the entry $\hat{z}_3$, $\sqrt{\Sigma_{33}}$ and the credible interval | 4.3.6 |
| Probability of staying below 28°C at 60 kW | the Gaussian tail $\Phi\big((T_{\max}-\hat{z}_3)/\sqrt{\Sigma_{33}}\big)$ | 4.3.7 |
| Largest set point at an assurance level | the tail at the node inverted, $T_{\max}-Q_2\hat{z}_1-\Phi^{-1}(\kappa)\sqrt{\Sigma_{33}}$ | 4.3.8 |
| *Further quantities reused later* | | |
| Spread of the conductance | $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ with $\mathbf{g}=(-1/\hat{z}_1^2,0,0)^{\mathsf{T}}$, and the exact push-forward | 4.3.6 |
| The entry that drives the temperature | the gradient and the variance budget | 4.3.2 |
| Effect of another reading | the information matrix plus the reading's $\mathbf{j}^{\mathsf{T}}\mathbf{j}$ | 4.6.1 |

The table is one inference followed by a column of cheap queries on its result, the division of Section 4.1.2. Two kinds of question fall outside it: a question that compares two models, which needs the evidence of each (Section 4.2.2), and the arrival of new evidence, which changes the posterior itself (Section 4.6.4).

Table 4.16 lists the five mathematical objects of the chapter that recur throughout the book, with the form of each and its purpose. The first three are the working objects when the posterior is well represented by a local Gaussian. The posterior distribution itself is the more fundamental object: it may be Gaussian, multimodal, discrete, a mixture or otherwise non-Gaussian, and its structure decides which computational method is appropriate.

Table 4.16: The five mathematical objects of an inference.

| Object | Form | Purpose |
|---|---|---|
| State estimate | $\hat{\mathbf{z}}$ | the most probable (or otherwise selected) central state |
| Posterior covariance | $\boldsymbol{\Sigma}$ | local uncertainty of the state and the dependence among its entries |
| Output gradient | $\mathbf{g} = \nabla f(\hat{\mathbf{z}})$ | connects state uncertainty to a derived quantity |
| Posterior distribution | $p(\mathbf{z} \mid \mathbf{y}, M)$ | the uncertainty about the state given the evidence |
| Model evidence | $p(\mathbf{y} \mid M)$ | normalizes the posterior and compares models |

Table 4.17 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 4.17: The terms of the anatomy of an inference and where each is defined. Units for the equipment hall.

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| Thermal resistance | $\rho = 1/G$, $T = T_c + Q\rho$ | K/kW | makes every row linear in the state | 4.1.4 |
| Physical rows | $r_1 = T_1 - 18 - 30\rho$, $r_2$ | K | conservation at each hall node | 4.1.5 |
| The question as one marginal | $\Pr(T_2 < T_{\max})$ from $\hat{z}_3$, $\Sigma_{33}$ | – | one entry's mean and variance | 4.1.6 |
| Model tolerance | $\sigma_p$; exact rows as $\sigma_p \to 0$ | K | model discrepancy of a physical row | 4.1.5 |
| Likelihood | $p(y_T \mid \mathbf{z})$, eq. (4.14) | $\mathrm{K}^{-1}$ | the reading as a function of the state | 4.2.1 |
| Residual vector | $\mathbf{r}(\mathbf{z})$, eq. (4.23) | – | four rows, each over its standard deviation | 4.2.3 |
| Objective | $\mathcal{L} = \frac{1}{2}\mathbf{r}^{\mathsf{T}}\mathbf{r}$ | – | minus log posterior | 4.2.3 |
| Jacobian | $\mathbf{J} = \partial\mathbf{r}/\partial\mathbf{z}$, $4 \times 3$ | mixed | one line per row, one column per entry | 4.2.3 |
| Information matrix | $\mathbf{J}^{\mathsf{T}}\mathbf{J} = \sum_k \mathbf{j}_k^{\mathsf{T}}\mathbf{j}_k = \tilde{\mathbf{J}}^{\mathsf{T}}\mathbf{W}\tilde{\mathbf{J}}$ | mixed | the Hessian when $\mathbf{C} = \mathbf{0}$ | 4.2.3 |
| General objective | $\mathcal{L} = \sum_k \mathcal{L}_k$, $\mathcal{L}_k = -\ln p_k$ | – | one term per row, any distribution | 4.2.7 |
| Count row, working form | $r_k = (\mu_k - n_k)/\sqrt{\mu_k}$, $\mathbf{j}_k = \sqrt{\mu_k}\,\nabla \ln \mu_k^{\mathsf{T}}$ | – | gradient and information, not the term | 4.2.7 |
| Hessian | $\mathbf{H} = \mathbf{J}^{\mathsf{T}}\mathbf{J} + \mathbf{C}$ | mixed | the information plus a named curvature term | 4.2.7 |
| Normal equations | $(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}$ | mixed | Gauss–Newton step to $\hat{\mathbf{z}}$ | 4.2.4 |
| Posterior covariance | $\mathbf{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ | mixed | spread is inverse information | 4.2.5 |
| Laplace approximation | $\mathcal{N}(\hat{\mathbf{z}}, \mathbf{H}^{-1})$ | – | local Gaussian posterior | 4.2.6 |
| Free direction | $\mathbf{v}$, $\mathbf{J}_{\mathrm{phys}}\mathbf{v} = \mathbf{0}$ | mixed | $\mathbf{\Sigma} \to \mathbf{v}\mathbf{v}^{\mathsf{T}}/5200$, rank one | 4.2.4 |
| Delta method | $\mathrm{Var}(f) \approx \mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}$ | varies | linearized propagation | 4.3.1 |
| Change of variables | $p_u(u) = p_x(h^{-1}(u))\,\lvert \mathrm{d}x/\mathrm{d}u \rvert$, eq. (4.55) | varies | density of a transformed quantity | 4.3.3 |
| Variance budget | eq. (4.54) | $\mathrm{K}^2$ | one term per entry plus cross terms | 4.3.2 |
| Threshold probability | $\Phi\big((T_{\max} - \hat{z}_3)/\sqrt{\Sigma_{33}}\big)$ | – | normal tail at the limit | 4.3.5 |
| Credible interval | $\hat{z}_3 \pm 1.960\sqrt{\Sigma_{33}}$ | °C | central 95 % posterior range | 4.3.4 |
| Scenario node | $T_2^*(r) = r + Q_2\rho$, eq. (4.63) | °C | the set point pinned at a value not yet applied | 4.3.8 |
| Back-off | $\Phi^{-1}(\kappa)\sqrt{\Sigma_{33}}$, eq. (4.66) | K | distance the set point stands inside the value that puts the mean on the limit, for the spread | 4.3.8 |
| Reduced state | $\mathbf{N}$, $\mathbf{J}_{\mathrm{red}} = \mathbf{J}_{\mathrm{rows}}\mathbf{N}$, $\mathbf{\Sigma}_{\mathrm{red}}$ | mixed | the free entries with exact physical rows | 4.2.8 |
| Prior predictive | $\mathcal{N}(22.8, 0.9014^2)$ | °C | reading predicted before use | 4.4.2 |
| Posterior predictive | $\mathcal{N}(23.077, 0.6504^2)$ | °C | next reading predicted | 4.4.4 |
| Evidence | $p(y_T \mid M)$, eq. (4.18) | $\mathrm{K}^{-1}$ | normalizer; compares models | 4.2.2 |
| Bayes factor | $B_{21} = p(y_T \mid M_2)/p(y_T \mid M_1)$ | – | evidence ratio of two models | 4.2.2 |
| Rank-one update | eq. (4.74) | – | adds one reading to a posterior | 4.6.4 |
| Preposterior covariance | $\mathbf{\Sigma}_a = (\mathbf{\Sigma}^{-1} + \mathbf{\Lambda}_a)^{-1}$, eq. (4.73) | mixed | ranks candidate readings | 4.6.1 |

## 4.8 Exercises

**Exercise 4.1** (Application: a complete inference)**.** A different equipment hall obeys $T = T_c + Q\rho$ with its supply air held at $T_c = 20\,^{\circ}\mathrm{C}$ and a test load $Q = 20\,\mathrm{kW}$. The prior on the resistance is $\rho \sim \mathcal{N}(0.20, 0.04^2)$ in K/kW, and a sensor with $\sigma_T = 0.4\,\mathrm{K}$ reads $y_T = 24.5\,^{\circ}\mathrm{C}$. The question is the temperature at a requested load of 40 kW. (a) Write the state $\mathbf{z} = (\rho, T_1, T_2)^{\mathsf{T}}$, with $T_1$ the test temperature and $T_2$ the temperature at 40 kW, the four normalized rows, and the $4 \times 3$ Jacobian at $\sigma_p = 0.01\,\mathrm{K}$. ($\mathbf{r} = ((T_1 - 20 - 20\rho)/\sigma_p,\ (T_2 - 20 - 40\rho)/\sigma_p,\ (T_1 - 24.5)/0.4,\ (\rho - 0.20)/0.04)^{\mathsf{T}}$; $\mathbf{J}$ has the lines $(-2000, 100, 0)$, $(-4000, 0, 100)$, $(0, 2.5, 0)$ and $(25, 0, 0)$.) (b) Take one Gauss–Newton step from the handbook state and compute $\boldsymbol{\Sigma}$ in the limit of exact rows. (At $\mathbf{z}^{(0)} = (0.20,\ 24.0,\ 28.0)^{\mathsf{T}}$ only the reading row is nonzero, $-1.25$, and $\mathbf{J}^{\mathsf{T}}\mathbf{r} = (0,\ -3.125,\ 0)^{\mathsf{T}}$; the normal equations give $\hat{\mathbf{z}} = (0.2200,\ 24.400,\ 28.800)^{\mathsf{T}}$, and a second step does not move it. The free direction is $\mathbf{v} = (1,\ 20,\ 40)^{\mathsf{T}}$, with $\mathbf{J}\mathbf{v} = (0,\ 0,\ 50,\ 25)^{\mathsf{T}}$ and information $50^2 + 25^2 = 3125$, so $\boldsymbol{\Sigma} \to \mathbf{v}\mathbf{v}^{\mathsf{T}}/3125$: $\Sigma_{11} = 3.2 \times 10^{-4}$, $\sqrt{\Sigma_{11}} = 0.01789\,\mathrm{K/kW}$; at $\sigma_p = 0.01\,\mathrm{K}$ the fit gives $\Sigma_{11} = 3.2016 \times 10^{-4}$ and a correlation of 0.9996 between $\rho$ and $T_1$.) (c) Read off the posterior mean and standard deviation of the temperature at 40 kW, and the probability that it stays below 30 °C. ($\hat{z}_3 = 28.8\,^{\circ}\mathrm{C}$ and $\sqrt{\Sigma_{33}} = \sqrt{1600/3125} = 0.7155\,\mathrm{K}$; $\Pr(T_2 < 30 \mid \mathbf{y}) = \Phi(1.2/0.7155) = \Phi(1.677) = 0.9532$, against $\Phi(2/1.6) = 0.8944$ under the prior alone.)

**Exercise 4.2** (Derivation: a variance budget with correlated entries)**.** For $T = T_c + Q\rho$ at the known load $Q = 50\,\mathrm{kW}$, the posterior covariance of $(T_c, \rho)$ is $\boldsymbol{\Sigma} = \begin{bmatrix} 0.25 & -0.003 \\ -0.003 & 0.0001 \end{bmatrix}$, in $\mathrm{K}^2$, K·K/kW and $(\mathrm{K/kW})^2$. (a) Derive the temperature variance from $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ and compute the standard deviation. ($\mathbf{g} = (1, 50)^{\mathsf{T}}$; $\mathrm{Var}(T) = 0.25 + 50^2 \times 0.0001 + 2 \times 50 \times (-0.003) = 0.25 + 0.25 - 0.30 = 0.20\,\mathrm{K}^2$, $\mathrm{sd}(T) = 0.4472\,\mathrm{K}$.) (b) State the result if the covariance term were ignored, and explain the difference. ($\sqrt{0.50} = 0.7071\,\mathrm{K}$, 58 % too large; the posterior correlation is $-0.003/\sqrt{0.25 \times 0.0001} = -0.6$, and with sensitivities of the same sign the negative cross term cancels part of the two diagonal terms: the evidence pins the combination $T_c + 50\rho$ better than either entry.)

**Exercise 4.3** (Derivation: a nonlinear derived quantity)**.** With $\rho \mid \mathbf{y} \sim \mathcal{N}(0.169231, 0.013868^2)$ from the chapter's example, derive the first-order approximation to the variance of the conductance $G = 1/\rho$, compute its standard deviation, and explain why the result is only an approximation. (With $G(\mathbf{z}) = 1/z_1$, $\mathbf{g} = (-1/\hat{z}_1^2,\ 0,\ 0)^{\mathsf{T}}$, so $\mathrm{Var}(G \mid \mathbf{y}) \approx \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} = \Sigma_{11}/\hat{z}_1^4$; $\hat{G} = 1/0.169231 = 5.909\,\mathrm{kW/K}$ and $\mathrm{sd}(G \mid \mathbf{y}) \approx 0.013868/0.169231^2 = 0.484\,\mathrm{kW/K}$. The map $1/\rho$ is nonlinear, so $G$ is not Gaussian: its exact 5–95 % band is $[5.207, 6.830]$, skewed, against the linearized $[5.113, 6.706]$; strictly, the moments of $1/\rho$ do not exist under a Gaussian $\rho$, which has positive density at zero, so positivity and the transformed distribution must be handled explicitly when exact conductance statements are needed.)

**Exercise 4.4** (Formulation: observability and the plant graph)**.** In $T = T_c + Q\rho$ suppose that both $T_c$ and $\rho$ are unknown, the discharge controller's set point not being trusted, and that there is a single temperature reading at one known load. (a) Write the observation Jacobian with respect to $(T_c, \rho)$ and state whether the reading can determine both unknowns without further information. ($y = T_c + Q\rho + \epsilon$, so $\mathbf{J} = [1 \quad Q]$, of rank one for two unknowns: the reading determines only the combination $T_c + Q\rho$, and every pair with the same combination explains it equally well.) (b) Name a reading or operating condition that resolves the ambiguity. (A second reading at a different known load $Q_2 \neq Q_1$ gives $\mathbf{J} = \begin{bmatrix} 1 & Q_1 \\ 1 & Q_2 \end{bmatrix}$, with determinant $Q_2 - Q_1 \neq 0$ and full rank; alternatively, a direct reading of the supply air, by a discharge sensor, a new node observed in the graph, removes the ambiguity.)

**Exercise 4.5** (Application: Bayesian evidence)**.** A model predicts $Y \sim \mathcal{N}(24, 1^2)$ in °C, and the reading $y = 25$°C arrives. (a) Compute the marginal likelihood $p(y \mid M)$. ($\frac{1}{\sqrt{2\pi}} \exp[-(25-24)^2/2] = 0.39894 \times 0.60653 = 0.2420\,\mathrm{K}^{-1}$.) (b) Explain why this value is not the probability that $M$ is correct. (It is a probability density in inverse temperature units, $2.420 \times 10^{-4}$ per millikelvin for the same reading, not $\Pr(M \mid y)$; a posterior model probability requires prior probabilities and the evidence of the alternative models under comparison.)

**Exercise 4.6** (Formulation: model validity and thermal storage)**.** The chapter assumed the steady-state balance $Q = G(T - T_c)$. The load now steps from 30 kW to 60 kW, and the engineer asks for the temperature five minutes later. (a) Explain why the steady-state equation may be insufficient. (It neglects the energy stored in the air, structure and equipment of the hall, so it gives only the temperature reached long after the change; five minutes may be short against the thermal time constant.) (b) Starting from an energy balance, formulate a model with thermal storage and name the new parameter and the condition it needs. (With an effective thermal capacitance $C_{\mathrm{th}}$ in kJ/K, $C_{\mathrm{th}}\,\mathrm{d}T/\mathrm{d}t = Q - G(T - T_c)$; the new parameter is $C_{\mathrm{th}}$ and an initial temperature $T(0)$ is required. For constant $Q$, $G$ and $T_c$ after the step, $T(t) = T_c + Q/G + [T(0) - T_c - Q/G]\,\mathrm{e}^{-Gt/C_{\mathrm{th}}}$, with time constant $C_{\mathrm{th}}/G$; the probability for the temperature at five minutes propagates the posterior of $G$, $C_{\mathrm{th}}$ and $T(0)$ through this time-dependent model rather than through the steady-state equation.)

**Exercise 4.7** (Application: the set point at another level)**.** With the posterior of Section 4.3.8, take the assurance level $\kappa = 0.90$. (a) Compute the largest set point $r_{90}$ and its back-off. ($\Phi^{-1}(0.90) = 1.2816$; the back-off is $1.2816 \times 0.8322 = 1.0665\,\mathrm{K}$, and equation (4.66) gives $r_{90} = 28 - 10.1538 - 1.0665 = 16.7797$°C.) (b) Compute the value of a second reading at the test condition that leaves $r_{90}$ unchanged. (The sure release is $1.2816 \times (0.8322 - 0.6399) = 0.2465\,\mathrm{K}$; equation (4.69) vanishes at $23.0769 + 0.2465/0.81806 = 23.3782$°C.)

# Part II

# The Closed-Loop Problem

# Chapter 5

# Actuators and Controllers as Graph Rows

## 5.1 The loop drawn in the graph

The commissioning engineer has twenty minutes of the discharge loop's trend open, at about 30 kW with the set point at $r = 18$°C. The means of the window read 22.8645°C on the hall sensor, 18.0032°C on the discharge sensor, 0.5635 on the valve's position feedback and 30.0500 kW on the load meter. The commissioning sheet takes those readings as exact and the coil's rating as true. It gives a hall conductance of 6.18 kW/K and a disturbance gain of 0.1619 K/kW, and, from the rated coil of 16 kW/K, its one forward line: lowering the set point to 17°C at 45 kW needs the valve at 0.938, so there is room. The trend's loop error is +0.003 K, so the sensor and the loop look fine.

One line of the sheet does not fit the others. Its own back-calculation of the coil from the same window gives 13.33 kW/K against the rated 16, and the sheet has no way to say which to believe. The sibling volume *Operations and Control by Derivation* reads its coils with nothing controlling the discharge, so that a warm discharge reading reports a weak coil. Here the valve moves. The discharge reading now reports what the controller did with it, and Part I's ideal loop is no longer enough. The question of this chapter is

*what closing a loop changes in the graph of a plant, and what the same move means in the two graphs.*

**Jargon: open loop, closed loop, actuator, stop, PI controller, integrator, loop error, disturbance, disturbance gain, process gain, controller's row, commissioning sheet, rating, reweighted draws, effective sample size**

A plant runs *open loop* when its actuator holds a position that nothing corrects, and *closed loop* when a controller moves the actuator to hold a reading at a set point. The *actuator* here is the chilled-water valve, with position $u$ between its two *stops*, 0 shut and 1 fully open. A *PI controller* adds to its command a term proportional to the *loop error* $e = y_c - r$, the reading less the set point, and an *integrator*, the running sum of that error. A *disturbance* is an input the controller does not set, here the load; the *disturbance gain* is the steady change of an output per unit of it, and the *process gain* the same per unit of the actuator. The *controller's row* (the framework's name for the controller in the graph) is the relation the controller imposes at steady state, written as a row of the model. The *commissioning sheet* is the engineer's page of forward calculations from single values; the *rating* is the coil's conductance from its data sheet. *Reweighted draws* are draws from the Laplace normal with weights that restore the exact posterior of the rows; the *effective sample size* says how many unweighted draws they are worth.

**Key idea**

Closing a loop does not change the number of unknowns: it swaps which entry is unknown. The set point is pinned instead of the valve, the valve is freed, and the controller's row holds the supply air. The same what-if then puts its band on the actuator instead of on the temperature, and at steady state, in window means, the loop's own reading stops informing the fit.

### 5.1.1 A controller as a row

Chapter 4 let the discharge controller hold the supply air at its set point and drew no valve. That was a statement about the graph: $T_c$ was a pinned boundary. To ask what the controller is doing, it has to appear in the graph with the same standing as the physics. The hall's controller is sampled every 30 s. At step $k$ it forms

$$e_k = y_{c,k} - r, \qquad c_k = u_0 + K_p\, e_k + I_k, \qquad I_{k+1} = I_k + K_i\, \Delta t\, e_k, \tag{5.1}$$

with the command $c_k$ going to the valve, $u_0 = 0.5$, $K_p = 0.10$ per kelvin, $K_i = 0.002$ per kelvin second and $\Delta t = 30$ s. The integrator holds while the command is at a stop and the error pushes it further, $c_k \geq 1$ with $e_k > 0$ or $c_k \leq 0$ with $e_k < 0$ (conditional integration), so it does not wind up. The controller is direct acting: its output rises with the measurement, and warm air opens the valve. How it reaches steady state, and whether it does, is the subject of Chapter 8. What it does when it gets there is one line of equation (5.1). At steady state the integrator does not move on average, so the mean of $e_k$ is zero and the mean of the reading is the set point. Writing the reading as the supply air plus a sensor bias $b$ and an error of zero mean, the steady controller imposes

$$T_c + b = r. \tag{5.2}$$

This is a row like any other. It holds a reading, not a temperature, and it holds only while the valve is off its stops, since the integrator needs the valve free to act on it (Section 5.3.5). The row is a law taken as given here; Chapter 7 states the problem it belongs to, and Chapters 12 to 14 derive where it comes from.

> **Principle 5.1.** *A controller with integral action adds one row at steady state, the reading held at the set point; its integrator adds one entry and one row that cancel. An actuator it moves is free in the graph; an actuator nothing moves is pinned.*

### 5.1.2 The closed-loop hall (step 0)

Figure 5.1 draws the hall of Part I with its discharge loop. The load enters the hall air on the left. Heat passes along the air path to the coil and supply air, and from the coil to the chilled water on the right, pinned. The lower row is the loop: the discharge sensor's reading goes to the PI controller, which compares it with the set point and commands the valve, and the valve sets how much of the coil's conductance is in use. Green boxes hold free quantities, gray boxes pinned ones, and dashed arrows lead to readings.

**The physical rows.** The state of the graph has six entries,

$$\mathbf{z}_{\text{full}} = (T,\ T_c,\ u,\ Q,\ g,\ g_w)^{\mathsf{T}}, \tag{5.3}$$

the hall air and the supply air in °C, the valve dimensionless, the load in kW, and the two log-factors $g = \ln(G/G_{\text{ref}})$ and $g_w = \ln(W/W_{\text{rated}})$, with $G_{\text{ref}} = 6$ kW/K and $W_{\text{rated}} = 16$ kW/K pinned. The integrator's state is left out, for the reason counted in Section 5.1.4. At steady

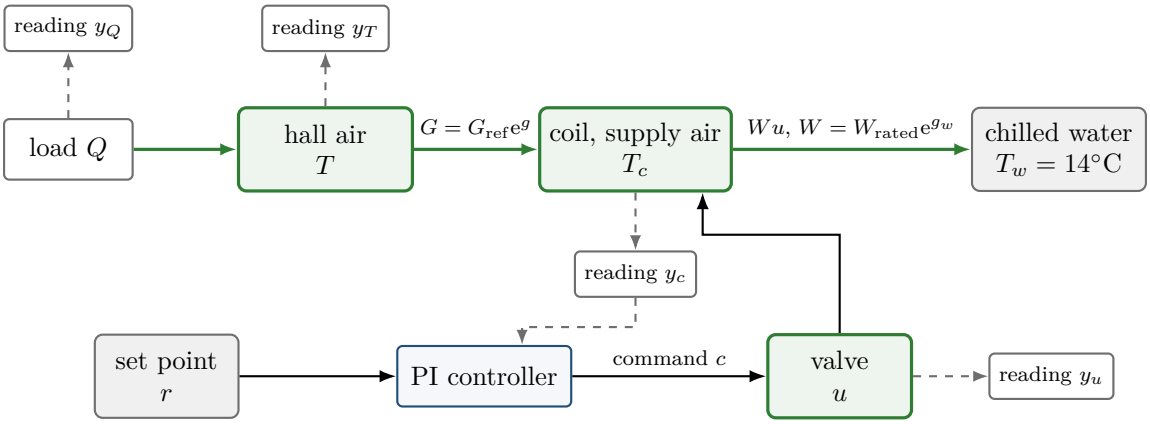



Figure 5.1: The closed-loop hall. Free (green): the hall air $T$, the supply air $T_c$ and the valve $u$; the load, the air path's log-conductance $g$ and the coil's log-derate $g_w$ are free entries carried on the paths. Pinned (gray): the chilled water and the set point. Dashed: the four readings; the discharge reading also feeds the controller. The controller is the only part of the figure that is not physics.

state the hall balance is $Q = G(T - T_c)$ and the coil balance is $Q = Wu\,(T_c - T_w)$. As rows that vanish when they hold, each with its tolerance,

$$r_{\text{h}} = \frac{T - T_c - (Q/G_{\text{ref}})\,\mathrm{e}^{-g}}{\sigma_p}, \qquad r_{\text{w}} = \frac{u - Q\,\mathrm{e}^{-g_w}/(W_{\text{rated}}(T_c - T_w))}{0.01\,\sigma_p}, \tag{5.4}$$

the second tolerance scaled because the valve is dimensionless and runs from 0 to 1.

**Modeling choices at the graph.** The chapter reads one steady operating point, the means of steps 40 to 79 of the record, twenty minutes at about 30 kW with $r = 18°\text{C}$: this makes every row algebraic and one window one state, and it fails in the minutes after a step, which Chapter 8 takes up. The coil is one conductance proportional to the valve position, which is what gives the valve one row, and it fails if the valve's characteristic is not linear.

The third choice is the one the rest of the chapter rests on. **The discharge sensor is taken as true**: $b = 0$ in equation (5.2), so that in the closed graph $T_c = r$ (the *sensor-true choice*). What it buys is a state of three entries in place of four. How it fails is that the fit then has no entry that can absorb a misreading of the supply air, and every quantity computed through $T_c$ inherits it; the checks of Section 5.4.4 test the answers, and Chapter 6 takes the choice up.

### 5.1.3 The control engineer's questions (step 1)

Table 5.1 lists the five questions asked of the window. The first column is the question in the engineer's words; the second the general form of Chapter 2 it takes; the third its mathematical form on the hall, with a star on the plant quantities at a scenario node; the fourth its category; the fifth the computation of Section 2.3.6 that answers it; and the last the subsection where it is answered. Each what-if is a scenario node in the sense of Section 4.3.8 and is asked of both graphs: the open node pins the valve, and the closed node pins the set point and leaves the valve to the controller's row.

### 5.1.4 The two graphs and the evidence as rows (step 2)

**The readings and the priors.** The evidence is the same in both graphs. Each reading is a window mean with its sensor's standard deviation divided by $\sqrt{40}$: $y_T = 22.8645°\text{C}$ (sd 0.0791),

Table 5.1: The questions asked of the closed-loop hall's window. $\mathbf{y}$ is the window's readings and the priors; a star marks a value at a scenario node.

| Control engineer's question | General form | Mathematical form | Category | Computation | Answered in |
|---|---|---|---|---|---|
| What does closing the loop change in the graph? | moments, (2.10) | the free and pinned entries of each graph; the columns of $\mathbf{J}$ | estimation | posterior | 5.3.1 |
| What is the plant now? | moments of derived quantities, (2.10) | $p(G, K_Q, W \mid \mathbf{y})$ | estimation of a derived quantity | posterior | 5.3.2 |
| Lower the set point 1 K: what moves, closed loop and open? | moments at a scenario node, (2.10) | $p(T^*, T_c^*, u^* \mid \mathbf{y})$ at $r^* = 17°\mathrm{C}$, against $u$ pinned | prediction | posterior | 5.3.3 |
| A load step to 45 kW: what moves, closed loop and open? | the same, (2.10) | the same at $Q^* = 45\,\mathrm{kW}$ | prediction | posterior | 5.3.4 |
| Will the valve reach its stop at 17°C and 45 kW? | threshold at a scenario node, (2.12) | $\Pr(u^* \geq 1 \mid \mathbf{y})$ | prediction under a threshold | posterior | 5.3.5 |

$y_c = 18.0032°\mathrm{C}$ (sd 0.0158), $y_u = 0.5635$ (sd 0.0016) and $y_Q = 30.0500\,\mathrm{kW}$ (sd 0.1581). Two priors are attached: the series' handbook on the air path, $g \sim \mathcal{N}(\ln(7.5/6),\ 0.3^2)$, and the coil's rating, $g_w \sim \mathcal{N}(0,\ 0.15^2)$. The load has none, since the meter reads it. Dividing by $\sqrt{40}$ takes the forty sensor errors as independent, which is what lets one window stand as one row per sensor, and that part is fair. What fails is the rest of the drawing, $T_c = r$ and static rows on window means, and a study over fresh records measures it (Section 5.6.5). The priors are taken as independent of each other, which keeps their rows in separate columns.

**The two graphs.** Figure 5.2 draws the same six entries twice, the open graph on the left and the closed on the right. Round boxes are entries, green when free and gray when pinned; small black squares are the rows, with a line to each entry the row touches; dashed boxes are the readings. The priors on $g$ and $g_w$ are below them. Table 5.2 says the same in words: each row of the table is a quantity and the two columns its role in each graph.

Table 5.2: The entries and rows of each graph.

| Quantity | Open graph | Closed graph |
|---|---|---|
| $T$ | free, read by $y_T$ | free, read by $y_T$ |
| $T_c$ | free, read by $y_c$ | free, fixed by the controller's row |
| $u$ | pinned at the value of $y_u$, whose spread drops | free, read by $y_u$ |
| $r$ | not in the graph | pinned |
| $Q$, $g$, $g_w$ | free; meter and priors | free; meter and priors |
| $y_c$ | a row on $T_c$ | no entry: a zero row |
| Physical and controller's rows | 2: hall, coil | 3: hall, coil, controller |
| Free entries | 5 | 6 |
| Left free once the rows hold | 3: $(Q, g, g_w)$; 4 with $u$ freed and $y_u$ a row | 3: $(Q, g, g_w)$; with the integrator, $7 - 4 = 3$ |

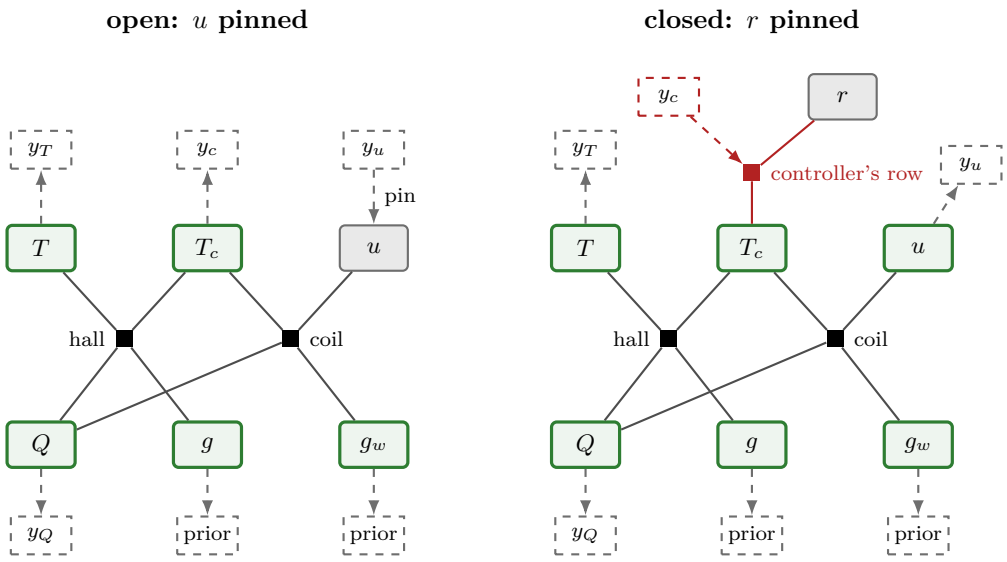



Figure 5.2: The two graphs of one hall. Open: the valve is pinned at the value its reading gives, and the discharge reading is a row on $T_c$. Closed: the set point is pinned, the valve is free and read, and the controller's row (red) holds the supply air; the discharge reading goes into that row and touches no entry of the state, by the sensor-true choice. Both graphs leave three entries free once their rows hold.

The count is the first thing to read. The open graph has two rows on five free entries, and three are left. The closed graph has three rows on six free entries, and three are left. Closing the loop swaps which entry is an unknown: the valve is freed, and the row that frees it is the one that fixes the supply air. Two things stand behind that count. The open graph's pin drops the spread of $y_u$: freed, with $y_u$ a row, it has four entries left, and the spread of $g_w$ goes from 0.00657 to 0.00714. And the closed graph's integrator adds one entry $I$ and one row, $u = u_0 + K_p e + I$, a net of zero: seven entries less four rows is still three. The reduced state of Section 4.2.8 is therefore the same three entries in both graphs,

$$\mathbf{z} = (Q,\ g,\ g_w)^{\mathsf{T}}, \tag{5.5}$$

in kW and two dimensionless entries. What differs is the map from it to the dependent entries. On the closed graph, with the sensor-true choice,

$$T_c = r, \qquad T = r + \frac{Q}{G_{\text{ref}}\, \mathrm{e}^{g}}, \qquad u = \frac{Q}{W_{\text{rated}}\, \mathrm{e}^{g_w}\,(r - T_w)}, \tag{5.6}$$

and on the open graph, with the valve pinned at $u$,

$$T_c = T_w + \frac{Q}{W_{\text{rated}}\, \mathrm{e}^{g_w}\, u}, \qquad T = T_c + \frac{Q}{G_{\text{ref}}\, \mathrm{e}^{g}}. \tag{5.7}$$

**Which row acts on which entry.** On the reduced state a reading sees the entries through the row of $\mathbf{N}_{\mathrm{b}}$ of what it reads, equation (4.47). With $*$ for an entry that can be nonzero and columns $Q, g, g_w$,

$$\mathbf{J}_{\text{closed}} = \begin{bmatrix} * & * & 0 \\ * & 0 & * \\ * & 0 & 0 \\ 0 & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{bmatrix} \begin{matrix} y_T \\ y_u \\ y_Q \\ y_c \\ \text{prior } g \\ \text{prior } g_w \end{matrix}, \qquad \mathbf{J}_{\text{open}} = \begin{bmatrix} * & * & * \\ * & 0 & * \\ * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{bmatrix} \begin{matrix} y_T \\ y_c \\ y_Q \\ \text{prior } g \\ \text{prior } g_w \end{matrix}. \tag{5.8}$$

Two facts are read here before any number. In the closed graph the discharge reading's line is all zeros: by equation (5.6) the supply air has a zero gradient on the state, so $y_c$ has no entry to act on. And the readings that see the coil's log-derate differ: the valve alone in the closed graph, the discharge sensor and the hall sensor in the open one, since $T$ there contains $T_c$.

## 5.2 The posterior of the state (step 3)

All five questions are read from one posterior, that of the reduced state under the closed graph's rows, because the loop in the record is closed. The open graph's fit of the same window is computed beside it.

### 5.2.1 The rows and the normal equations

The rows on the reduced state are the evidence rows at the point of the set, equation (4.46), with the dependent entries of equation (5.6). Each is divided by its standard deviation:

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T(\mathbf{z}) - y_T)/0.0791 \\ (u(\mathbf{z}) - y_u)/0.0016 \\ (Q - y_Q)/0.1581 \\ (g - \ln(7.5/6))/0.3 \\ g_w/0.15 \end{bmatrix}, \qquad \mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z}). \tag{5.9}$$

The discharge reading's row is left out of the five because it has no entry: it is the constant $(y_c - r)/0.0158$, a zero line of the Jacobian that adds nothing to the gradient or to the information. The reduced Jacobian is $\mathbf{J} = \mathbf{J}_{\text{rows}}\mathbf{N}$ of equation (4.47), and the rows of $\mathbf{N}_{\text{b}}$ are the gradients of equation (5.6),

$$\frac{\partial T}{\partial \mathbf{z}} = \Big(\frac{1}{G},\ -\frac{Q}{G},\ 0\Big), \quad \frac{\partial u}{\partial \mathbf{z}} = \Big(\frac{u}{Q},\ 0,\ -u\Big), \quad \frac{\partial T_c}{\partial \mathbf{z}} = (0,\ 0,\ 0), \tag{5.10}$$

with $G = G_{\text{ref}}\mathrm{e}^{g}$. The last line is the zero row of the discharge reading, now as an object of the fit. Each Gauss–Newton step solves the normal equations of equation (4.29), $(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}(\mathbf{z})$, with $\mathbf{J}$ and $\mathbf{r}$ at the current iterate, and starts from the meter's load, the handbook's air path and the rated coil. With normalized rows the information matrix is $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$; with unnormalized rows it is $\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$. The reported covariance is the inverse of the full Hessian at the minimum, equation (4.38), with the curvature term of the nonlinear rows included.

### 5.2.2 The fit of the window

Table 5.3 runs the normal equations. Each row is an iterate, and the columns give the three entries, the objective and the norm of its gradient. The first step moves both log-factors most of the way, and the objective falls from 1823.0166 to 0.9456 by the third.

**The Jacobian and the information.** At the MAP state, with columns $Q, g, g_w$ and rows in the order of equation (5.9),

$$\mathbf{J} = \begin{bmatrix} 2.0459 & -61.4967 & 0 \\ 11.85469 & 0 & -356.33437 \\ 6.32456 & 0 & 0 \\ 0 & 3.33333 & 0 \\ 0 & 0 & 6.66667 \end{bmatrix}, \quad \mathbf{J}^{\mathsf{T}}\mathbf{J} = \begin{bmatrix} 184.7194 & -125.81608 & -4224.23376 \\ -125.81608 & 3792.95523 & 0 \\ -4224.23376 & 0 & 127018.62742 \end{bmatrix}, \tag{5.11}$$

the first column of $\mathbf{J}$ in $\mathrm{kW}^{-1}$ and the others dimensionless, with the zeros where the pattern (5.8) puts them. The valve's line is the strongest in the matrix. The zero in the $(g, g_w)$ place of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ says that no row touches both, so the air path and the coil are tied only through the load.

Table 5.3: Gauss–Newton iterations of the closed fit on the reduced state, from the meter's load, the handbook's $g$ and the rated coil.

| $k$ [–] | $Q$ [kW] | $g$ [–] | $g_w$ [–] | $\mathcal{L}$ [–] | $\lVert\mathbf{J}^{\mathsf{T}}\mathbf{r}\rVert$ [–] |
|---|---|---|---|---|---|
| 0 | 30.0500 | 0.22314 | 0.00000 | 1823.0166 | $1.77 \times 10^{4}$ |
| 1 | 30.0593 | 0.01027 | −0.19961 | 21.8258 | $2.30 \times 10^{3}$ |
| 2 | 30.0585 | 0.02979 | −0.18215 | 0.9473 | $1.98 \times 10^{1}$ |
| 3 | 30.0585 | 0.02999 | −0.18200 | 0.9456 | $2.85 \times 10^{-3}$ |
| 4 | 30.0585 | 0.02999 | −0.18200 | 0.9456 | $3.42 \times 10^{-7}$ |
| 5 | 30.0585 | 0.02999 | −0.18200 | 0.9456 | $1.10 \times 10^{-9}$ |

**The covariance and the MAP state.** The inverse Hessian is

$$\boldsymbol{\Sigma} = \begin{bmatrix} 2.495 \times 10^{-2} & 8.278 \times 10^{-4} & 8.299 \times 10^{-4} \\ 8.278 \times 10^{-4} & 2.913 \times 10^{-4} & 2.753 \times 10^{-5} \\ 8.299 \times 10^{-4} & 2.753 \times 10^{-5} & 3.547 \times 10^{-5} \end{bmatrix}, \tag{5.12}$$

in $\mathrm{kW}^2$, kW and dimensionless units, and the MAP state with its spreads is

$$\hat{Q} = 30.0585\,\mathrm{kW}\ (0.1580), \qquad \hat{g} = 0.02999\ (0.01707), \qquad \hat{g}_w = -0.18200\ (0.00596). \tag{5.13}$$

The correlations are 0.3070 for $Q$ with $g$, 0.8821 for $Q$ with $g_w$ and 0.2708 for $g$ with $g_w$. Their signs come from the lines of $\mathbf{J}$. The valve's line has opposite signs in the columns of $Q$ and $g_w$: with the reading held, a larger load is explained by a stronger coil. The hall sensor's line does the same for $Q$ and $g$, more weakly because the hall sensor is the noisier reading. The high 0.8821 says that the coil's spread is mostly the load's. At the minimum the normalized rows are −0.035 for $y_T$, −0.023 for $y_u$, 0.054 for $y_Q$, −0.644 for the prior on $g$ and −1.213 for the prior on $g_w$: the three readings are met almost exactly, and the two priors carry the misfit.

### 5.2.3 The open fit of the same window

The open graph pins the valve at its reading, makes the discharge reading a row on $T_c$ by equation (5.7), and keeps the other rows. Table 5.4 sets the two fits side by side; each row is an entry or the misfit, and the columns are the two graphs.

Table 5.4: The closed and open fits of one window, with posterior spreads $\sqrt{\Sigma_{ii}}$ in parentheses. $2\mathcal{L}$ is on two degrees of freedom.

| Entry [unit] | Closed graph | Open graph |
|---|---|---|
| $Q$ [kW] | 30.0585 (0.1580) | 30.0585 (0.1580) |
| $g$ [–] | 0.02999 (0.01707) | 0.03058 (0.01738) |
| $g_w$ [–] | −0.18200 (0.00596) | −0.18277 (0.00657) |
| $2\mathcal{L}(\hat{\mathbf{z}})$ [–] | 1.8913 | 1.9015 |

## 5.3 The answers (step 4)

Every answer is read from $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ of Section 5.2.2 by two named routes. The *state route* is the delta method on the Laplace normal, a quantity at $\hat{\mathbf{z}}$ with variance $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$, equation (4.53); its value is the one reported. The *draws* are $10^6$ reweighted draws of the state, seed 5, each pushed through the exact forward equations, and they stand beside it, center against median and spread against spread.

### 5.3.1 What closing the loop changes

**Control engineer's question.** What does closing the loop change in the graph?

**General form.** An estimation question on an entry asks for its posterior moments, equation (2.10),

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad \mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y}) = \int \big(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]\big)^2\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

**Specialized form.** Equation (2.10) with $f(\mathbf{z}) = z_i$ and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal, taken under each graph's rows in turn. An entry's mean and variance are $\hat{z}_i$ and $\Sigma_{ii}$, and what a reading contributes to them is its outer product in $\mathbf{J}^\mathsf{T}\mathbf{J}$, equation (3.27). For the discharge reading in the closed graph that line is zero by equation (5.10), so its outer product is the zero matrix.

**Building blocks.** The patterns (5.8), the gradient (5.10) and the fits of Table 5.4.

**The answer on the hall.** Three things change. The count does not: three entries are left in each graph (Table 5.2), and the valve and the supply air trade places. The route of the information changes: the coil is read by the valve in the closed graph and by the discharge sensor in the open one. And at steady state, in window means, the discharge reading stops informing anything: in the closed graph it has a zero row, information 0, and its residual, $+0.0032\,\mathrm{K}$, is the loop's error. *The loop consumes its own reading.* The trend's flat error is therefore no evidence about the sensor in this graph, in either direction. The qualification matters: at a stop the controller's row is absent, and Section 5.4.3 reads $y_c$ there.

The entries barely move: $g_w = -0.18200$ (sd 0.00596) closed and $-0.18277$ (sd 0.00657) open. This agreement is not a confirmation. Both graphs take the sensor as true, the closed one by setting $b = 0$ and the open one by reading $y_c$ as $T_c$ with an unbiased error. They have the same hypothesis under them, and the record computes the same answer twice.

### 5.3.2 The plant now

**Control engineer's question.** What is the plant now?

**General form.** The moments of a derived quantity, equation (2.10) again, now with $f$ a function of several entries or of one entry through a nonlinear map.

**Specialized form.** Equation (2.10) with $f$ each of $G = G_{\mathrm{ref}}\mathrm{e}^{g}$, the disturbance gain $K_Q = 1/G$, $W = W_{\mathrm{rated}}\mathrm{e}^{g_w}$ and the process gain $K_u = \partial T_c/\partial u = -(T_c - T_w)/u$ of equation (5.7), read at $T_c = r$ and the $u$ of equation (5.6), and with $T$ and $u$ of that equation. The gradients on $(Q, g, g_w)$ are $(0, G, 0)$, $(0, -K_Q, 0)$, $(0, 0, W)$ and $(-K_u/Q, 0, K_u)$, and those of $T$ and $u$ are equation (5.10). By the delta method each variance is $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$.

**Building blocks.** The MAP state (5.13), the covariance (5.12) and these gradients.

**The plant's numbers.** Table 5.5 gives the values: each row is a quantity, the first two numeric columns the state route and the last two the draws. The coil is at $13.3377\,\mathrm{kW/K}$ with sd 0.0794, beside the sheet's back-calculated 13.3330 and against the rated 16. The fit has answered the sheet's unanswered line by weighing the two: the prior on the rating is about one of its own standard deviations wide of the answer, at $-1.213$, while the valve's reading is met to $-0.023$. $K_Q$ is what the load does to the hall air, the disturbance gain of Chapter 1. The

process gain, what the valve does to the supply air, first exists here, because the valve first exists here: $K_u = -7.0996\,\mathrm{K}$ per unit of valve position, sd 0.0199.

Table 5.5: The plant now, under the sensor-true choice: the state route, with the reweighted draws beside it.

| Quantity [unit] | State route | sd | Draws, median | Draws, sd |
|---|---|---|---|---|
| Hall air $T$ [°C] | 22.8617 | 0.0790 | 22.8605 | 0.0790 |
| Valve $u$ [–] | 0.5634 | 0.0016 | 0.5634 | 0.0016 |
| Conductance $G$ [kW/K] | 6.1827 | 0.1055 | 6.1845 | 0.1057 |
| Disturbance gain $K_Q$ [K/kW] | 0.1617 | 0.0028 | 0.1617 | 0.0028 |
| Coil $W$ [kW/K] | 13.3377 | 0.0794 | 13.3378 | 0.0794 |
| Process gain $K_u$ [K] | −7.0996 | 0.0199 | −7.0996 | 0.0199 |

### 5.3.3 The same move in the two graphs

**Control engineer's question.** Lower the set point 1 K: what moves, closed loop and open?

**General form.** A prediction at a scenario node asks for the moments of a quantity at a value not yet applied, equation (2.10) with $f$ evaluated at the node.

**Specialized form.** Equation (2.10) with $f$ each of $T_c^*$, $T^*$, $u^*$ at the present load, on the posterior of Section 5.2.2, with the node re-solved as in Section 4.3.8. The closed node pins $r^* = 17$°C and is equation (5.6) with $r^*$ in place of $r$. To compare like with like, the open node pins the valve at the closed node's mean, $u = 0.75122$, the position that would bring the supply air to 17°C at the MAP state, and it is equation (5.7). The gradients on $(Q, g, g_w)$ are

$$\begin{aligned} \text{closed: } & \mathbf{g}_{T_c^*} = \mathbf{0}, \quad \mathbf{g}_{T^*} = (0.16174,\ -4.86174,\ 0), \quad \mathbf{g}_{u^*} = (0.02499,\ 0,\ -0.75122), \\ \text{open: } & \mathbf{g}_{u^*} = \mathbf{0}, \quad \mathbf{g}_{T^*} = (0.26155,\ -4.86174,\ -3), \quad \mathbf{g}_{T_c^*} = (0.09981,\ 0,\ -3). \end{aligned} \tag{5.14}$$

The closed $\mathbf{g}_{u^*}$ and the open $\mathbf{g}_{T_c^*}$ are the same direction in the state: the open one is the closed one times $(T_c^* - T_w)/u^*$, by equations (5.6) and (5.7). One combination of $Q$ and $g_w$ carries the coil's uncertainty, and the graph decides where it lands.

**Building blocks.** The covariance (5.12) and the gradients (5.14).

**The hall's numbers.** Table 5.6 gives the two nodes: the columns are the three plant quantities in each graph, the first row the state route and the second the draws, with spreads in parentheses. Figure 5.3 draws the spreads of the supply air and the valve as bars, one panel for each quantity, one bar for each graph.

Table 5.6: The set point lowered to 17°C at the present load: the closed posterior through each map; mean (sd). The open node pins the valve at 0.7512, the closed node's mean.

| Route | Closed: $r^* = 17$°C pinned | | | Open: $u$ pinned | | |
|---|---|---|---|---|---|---|
| | $T_c^*$ [°C] | $T^*$ [°C] | $u^*$ [–] | $T_c^*$ [°C] | $T^*$ [°C] | $u^*$ [–] |
| State | 17 (0) | 21.8617 (0.0790) | 0.7512 (0.0021) | 17.0000 (0.0084) | 21.8617 (0.0794) | 0.7512 (0) |
| Draws | 17.0000 (0.0000) | 21.8605 (0.0790) | 0.7512 (0.0021) | 17.0000 (0.0084) | 21.8604 (0.0794) | 0.7512 (0.0000) |

*The loop moves the band from the temperature to the actuator.* Closed, $T_c^*$ has no spread and $u^*$ has it; open, the reverse. The table shows that at a stated strength, since it is one

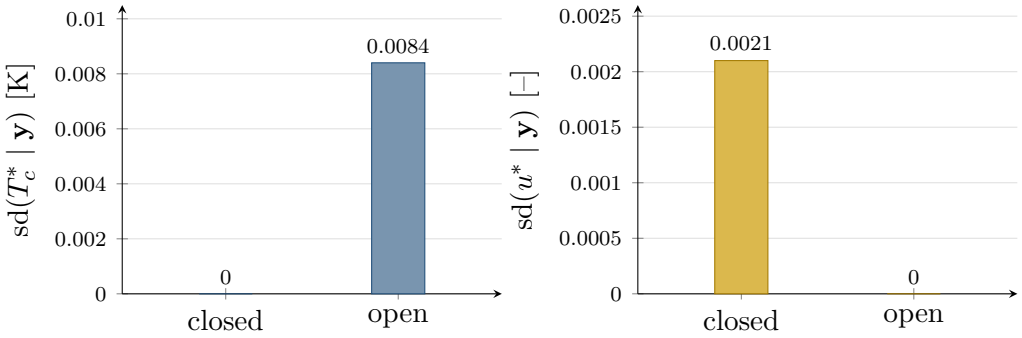



Figure 5.3: The same move, the set point lowered to 17°C at the present load: the spread of the supply air (left) and of the valve (right), the closed posterior through each map, from Table 5.6. The zero bars are the entry each graph pins or fixes.

posterior in two units. The open 0.0084 is the closed 0.0021 converted, $0.0021 \times 3/0.7512$; and $T_c^* = 17.0000$ and the shared 21.86°C are forced by the pin at the closed node's mean. That much is construction. The open graph's own fit, with its own $\mathbf{\Sigma}$, gives $T_c^* = 17.0023$°C with sd 0.0119 K and $T^* = 21.8612$°C with sd 0.0791 K; with $y_u$ a row instead of a pin, the spread of $T_c^*$ is 0.0145 K. The direction of the claim holds in all of them, and its size depends on the graph that was fitted. The zero spreads hold only as computed: the closed zero is the sensor-true choice with the controller's row exact, and the open zero is the pin. Neither says the supply air or the valve is known to that precision in the plant.

The band in the table is small on both maps for the same reason, the 0.8821 of Section 5.2.2. The variance of the open $T_c^*$ is 0.00025 from $Q$ and 0.00032 $\mathrm{K}^2$ from $g_w$, with a cross term of $-0.00050$; that of the closed $u^*$ is $1.559 \times 10^{-5}$ from $Q$ and $2.002 \times 10^{-5}$ from $g_w$, with a cross term of $-3.116 \times 10^{-5}$. The gradient has opposite signs on $Q$ and $g_w$ and their covariance is positive, so the cross term takes back most of the two diagonal terms.

### 5.3.4 A load step in the two graphs

**Control engineer's question.** A load step to 45 kW: what moves, closed loop and open?

**General form.** The same moments at a scenario node, equation (2.10).

**Specialized form.** Equation (2.10) with the load now pinned at the node, $Q^* = 45$ kW, in place of the entry $Q$ of the state, so the load's column of each gradient is zero. The closed node keeps $r = 18$°C and is equation (5.6). The open node leaves the valve at the present fit's $u = 0.56341$ and is equation (5.7). The gradients are in the two log-factors only: closed, $\mathbf{g}_{T^*} = (0,\ -Q^*/G,\ 0)$ and $\mathbf{g}_{u^*} = (0,\ 0,\ -u^*)$; open, $\mathbf{g}_{T_c^*} = (0,\ 0,\ -Q^*/(Wu))$ and $\mathbf{g}_{T^*} = (0,\ -Q^*/G,\ -Q^*/(Wu))$.

**Building blocks.** The covariance (5.12) and these gradients.

**The hall's numbers.** Table 5.7 lays out the two nodes as Table 5.6 did. Figure 5.4 draws the temperatures against the load pinned at the node, from 30 to 50 kW: the horizontal axis is the load, the vertical axis the temperature, the lines are the means by equations (5.6) and (5.7), and the shaded strips are ±1.96 of their standard deviations.

This row shows what the loop is for. The disturbance reaches the supply air only in the open graph: 19.988°C against 18°C, a difference of about 2 K, and the same difference is then carried to the hall air, 27.267°C against 25.278°C. In the closed graph the valve absorbs the

Table 5.7: A load step to 45 kW at $r = 18°C$: the closed posterior through each map; mean (sd). The open node leaves the valve at 0.5634.

| Route | Closed: $r = 18°C$ pinned | | | Open: $u$ pinned | | |
|---|---|---|---|---|---|---|
| | $T_c^*$ [°C] | $T^*$ [°C] | $u^*$ [–] | $T_c^*$ [°C] | $T^*$ [°C] | $u^*$ [–] |
| State | 18 (0) | 25.2784 (0.1242) | 0.8435 (0.0050) | 19.9883 (0.0357) | 27.2667 (0.1382) | 0.5634 (0) |
| Draws | 18.0000 (0.0000) | 25.2763 (0.1243) | 0.8435 (0.0050) | 19.9883 (0.0356) | 27.2645 (0.1383) | pinned |

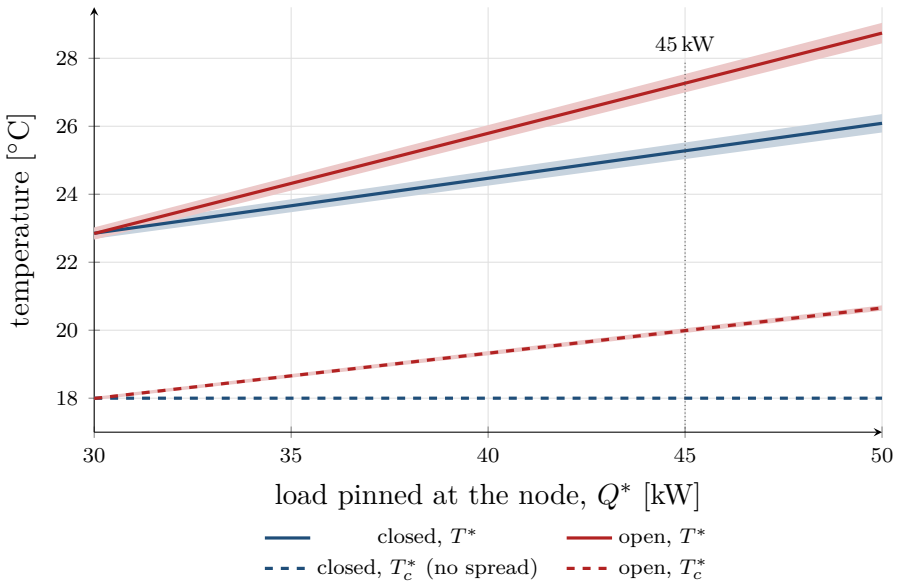



Figure 5.4: The supply air and the hall air against the load in each graph. Closed: $T_c^* = 18°C$ and $T^* = 18 + Q^*/G$ with $G = 6.1827\,\text{kW/K}$. Open: $T_c^* = 14 + Q^*/(Wu)$, with $Wu = 13.3377 \times 0.56341\,\text{kW/K}$, and $T^* = T_c^* + Q^*/G$. Strips are ±1.96 sd, the closed posterior through each map, the spreads of Table 5.7 scaled with $Q^*$. The dotted line marks the question's 45 kW.

load, from 0.5634 to 0.8435. The band moves with it, as in the set-point move. The open graph's own fit, with the valve left at its reading and its own $\boldsymbol{\Sigma}$, gives $T_c^* = 19.9926$°C with sd 0.0394 K and $T^* = 27.2667$°C with sd 0.1372 K, the same difference on a wider band.

### 5.3.5 The valve at its stop

**Control engineer's question.** Will the valve reach its stop at 17°C and 45 kW?

**General form.** A prediction under a threshold is a tail of a derived quantity, equation (2.12),

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \geq f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

**Specialized form.** Equation (2.12) with $f(\mathbf{z}) = u^*(\mathbf{z}) = Q^*/(W_{\text{rated}} \mathrm{e}^{g_w}(r^* - T_w))$ at the closed node of equation (5.6), $Q^*$ and $r^*$ pinned, $f^* = 1$ and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal. The assurance level $\kappa$ of that equation is not set; the probability is reported. Two things enter here. The value $u^*$ is what the controller's row would require of the valve. Above 1 it is not a position, since at a stop the actuator is pinned again, so the node answers whether the row can hold and no more; what happens beyond the stop is the subject of Chapter 15. And the normal tail is the tail of the delta method's normal, by equation (4.57),

$$\Pr(u^* \geq 1 \mid \mathbf{y}) = \Phi\left(\frac{u^*(\hat{\mathbf{z}}) - 1}{\sqrt{\mathbf{g}_{u^*}^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}_{u^*}}}\right), \qquad \mathbf{g}_{u^*} = (0,\ 0,\ -u^*). \tag{5.15}$$

**Building blocks.** $\hat{g}_w$ and $\Sigma_{33}$ of equations (5.13) and (5.12), the gradient of equation (5.15) and the normal distribution function $\Phi$.

**The hall's numbers.** At $r^* = 17$°C and $Q^* = 45$ kW the state route gives $u^* = 1.1246$, sd 0.0067; the draws give a median of 1.1246 and an sd of 0.0067. The tail is $\Phi(18.607)$, which prints as $\Pr(u^* \geq 1 \mid \mathbf{y}) = 1.0000$, and the draws print 1.0000. That figure is reported as what it is: a normal approximation far in its tail, under the sensor-true choice. It says the stop lies far below the fit's band and does not say the event is certain, since the sensor-true choice lies outside the band. Table 5.8 walks the set point down at 45 kW: each row is a set point, with the required valve position, its spread and the probability. Figure 5.5 draws the same thing: the horizontal axis is the set point at the node, the vertical axis the valve position; the green curve is the fit with its strip, the blue curve the rated coil, the red line the stop, and the dotted stems the record's stated truth, a check marker for Section 5.4.3.

Table 5.8: The valve position the controller's row requires at 45 kW, on the state route, under the sensor-true choice.

| $r^*$ [°C] | $u^*$ [–] | sd($u^*$) [–] | $\Pr(u^* \geq 1 \mid \mathbf{y})$ [–] |
|---|---|---|---|
| 18.00 | 0.8435 | 0.0050 | 0.0000 |
| 17.75 | 0.8997 | 0.0054 | 0.0000 |
| 17.50 | 0.9640 | 0.0057 | 0.0000 |
| 17.25 | 1.0381 | 0.0062 | 1.0000 |
| 17.00 | 1.1246 | 0.0067 | 1.0000 |

*The rated coil's* 0.938 *is overturned.* At the rated coil the valve has room at 17°C; on the fit it has none, and it crosses the stop between 17.50 and 17.25°C. The fit's center is the sheet's own back-calculated figure, 1.125; what the fit adds is that it has weighed this figure against the rating, and the spread that goes with it. Every number here carries the sensor-true choice.

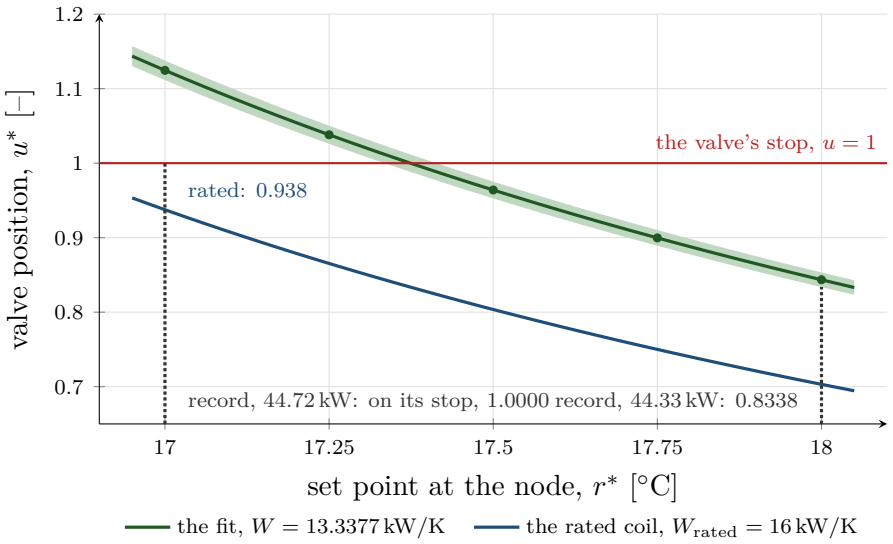



Figure 5.5: The valve position the controller's row requires at 45 kW. The fit is $u^* = 45/(W(r^* - 14))$ with its strip of ±1.96 sd (the spread is 0.00596 of $u^*$), and its points are those of Table 5.8; the rated coil is the same curve at $W = 16$ kW/K. The stems mark the record's stated truth where it ran near 45 kW with these set points, at 44.33 and 44.72 kW; the truth is not drawn between them.

## 5.4 The checks (step 5)

The answers rest on a limit, on a normal approximation, on a misfit that is small and on the choice of step 0. Each is tested here, and the last test fails.

### 5.4.1 The limit of exact rows

The fit runs on the reduced state, which equation (4.48) makes the limit $\sigma_p \to 0$ of the full state. The check fits the full state with the physical rows (5.4) at a finite tolerance, with $T_c$ held at $r$ by the sensor-true choice, so that the entries are $(T, u, Q, g, g_w)$. The covariance is the inverse of the numeric Hessian, the curvature term of the nonlinear rows included, since it does not vanish in the limit. Table 5.9 gives the result: each row is a tolerance and the columns the two log-factors with their spreads. At 0.1 the air path's spread is wider by the discrepancy the rows are allowed, and at 0.001 the fit is the reduced state to the digits printed. Dropping the curvature term altogether, in the Gauss–Newton form on the reduced state, gives spreads of 0.15799, 0.01706 and 0.00596 against the reported 0.15797, 0.01707 and 0.00596.

Table 5.9: The limit of exact physical rows on the full state of the closed graph. The valve row's tolerance is $0.01\,\sigma_p$.

| Tolerance $\sigma_p$ [K] | $g$ [–] | sd($g$) [–] | $g_w$ [–] | sd($g_w$) [–] |
|---|---|---|---|---|
| 0.1 | 0.03089 | 0.02668 | −0.18197 | 0.00621 |
| 0.01 | 0.03000 | 0.01719 | −0.18200 | 0.00596 |
| 0.001 | 0.02999 | 0.01707 | −0.18200 | 0.00596 |
| reduced state | 0.02999 | 0.01707 | −0.18200 | 0.00596 |

### 5.4.2 The state route against the draws

The draws of Section 5.3 are $N = 10^6$ states from the Laplace normal, seed 5, weighted to the exact posterior of the rows (5.9), with an effective sample size of 998 990: the Laplace normal is close to that posterior. Table 5.10 sets the values at the nodes of step 4 beside them, with $G$ from Table 5.5: each row is a quantity at its node, the first two numeric columns the state route and the rest the draws. The Monte Carlo standard error of a draws mean is its sd over the root of the effective size, about one thousandth of the sd.

Table 5.10: The state route against the reweighted draws ($N = 10^6$, seed 5, effective size 998 990). Units as in the tables of step 4.

| Quantity at its node | State | sd | Draws, median | mean | sd |
|---|---|---|---|---|---|
| Closed $T^*$, 17°C, present load [°C] | 21.8617 | 0.0790 | 21.8605 | 21.8605 | 0.0790 |
| Closed $u^*$, 17°C, present load [–] | 0.7512 | 0.0021 | 0.7512 | 0.7512 | 0.0021 |
| Open $T_c^*$, valve at 0.7512 [°C] | 17.0000 | 0.0084 | 17.0000 | 17.0000 | 0.0084 |
| Open $T^*$, valve at 0.7512 [°C] | 21.8617 | 0.0794 | 21.8604 | 21.8605 | 0.0794 |
| Closed $T^*$, 18°C, 45 kW [°C] | 25.2784 | 0.1242 | 25.2763 | 25.2767 | 0.1243 |
| Closed $u^*$, 18°C, 45 kW [–] | 0.8435 | 0.0050 | 0.8435 | 0.8435 | 0.0050 |
| Open $T_c^*$, valve at 0.5634, 45 kW [°C] | 19.9883 | 0.0357 | 19.9883 | 19.9885 | 0.0356 |
| Open $T^*$, valve at 0.5634, 45 kW [°C] | 27.2667 | 0.1382 | 27.2645 | 27.2651 | 0.1383 |
| Closed $u^*$, 17°C, 45 kW [–] | 1.1246 | 0.0067 | 1.1246 | 1.1247 | 0.0067 |
| $G$, the plant now [kW/K] | 6.1827 | 0.1055 | 6.1845 | 6.1859 | 0.1057 |
| $\Pr(u^* \geq 1 \mid \mathbf{y})$, 17°C, 45 kW [–] | 1.0000 | – | – | 1.0000 | – |

The largest gaps are on the air path: 0.0018 kW/K on $G$ and 0.0022 K on the open $T^*$ at 45 kW. They exceed the Monte Carlo error, so they are a real difference between the MAP state and the center of a posterior that is not exactly normal in the log-conductance, and they are small against the spreads, 0.1055 and 0.1382. The check on the tail is weaker than the others: at $\Phi(18.607)$ neither route resolves anything but 1.0000, and the gap prints as 0.0000 with a standard error that prints as 0.0000.

### 5.4.3 The record's own later step

The record continues past the window. Its load steps toward 45 kW, and with the loop in, over its steps 140 to 159, the load is 44.33 kW and the truth holds $T = 25.330$°C and $u = 0.8338$. The closed node at 45 kW reads 25.278°C and 0.8435; at the record's own load it reads 25.169°C and 0.8308. Then the set point steps to 17°C. Over steps 200 to 239 the load is 44.72 kW, where the controller's row would require $u^* = 1.1176$, and the truth has the valve on its stop, $u = 1.0000$. The integrator holds, by its conditional clause, so the command stays near 1.032, and the discharge reading stays at 17.460°C against $r = 17$°C: the loop has lost its set point, and with the row absent the reading speaks again. That agrees with the answer of Section 5.3.5, and overturns the rated 0.938 a second time. It agrees about the event; it does not vouch for the coil, as the next check shows.

### 5.4.4 The misfit and the truth

**What the misfit can see.** Under the model, twice the objective at the minimum is a chi-square on the rows less the entries, here $5 - 3 = 2$ degrees of freedom. The value is $2\mathcal{L}(\hat{\mathbf{z}}) = 1.8913$, and $\Pr(\chi^2_2 \geq 1.8913) = 0.388$. The open fit gives 1.9015 and 0.3865. That says little. Three readings meet three entries, so the readings are met nearly exactly by construction, and the two degrees of freedom are the priors': of the 1.8913, the priors contribute 1.8867 and the readings 0.0046. Across 200 fresh records of the same plant and the same truth (Section 5.6.5)

the value has a median of 1.8479 and a maximum of 2.3632, and none exceeded the 5.99 level of $\chi^2_2$. The test is blind to what follows.

**The truth row.** The record is generated from a stated truth, and Table 5.11 compares the fit with it: each row is a quantity, the columns the fit, its posterior spread, the truth and, for the entries of the state, the truth's distance from the fit in the fit's own standard deviations. Figure 5.6 draws those distances as bars on the left, with the dashed lines at ±1.96; on the right it draws the chi-square density on two degrees of freedom, with the misfit marked and the area beyond it shaded.

Table 5.11: The closed fit against the record's stated truth over the same steps. The gap in $T_c$ is reported as the record shows it; Chapter 6 asks its cause.

| Quantity [unit] | Fit | sd | Truth | Distance [sd] |
|---|---|---|---|---|
| $Q$ [kW] | 30.05851 | 0.15797 | 30.25634 | +1.25 |
| $g$ [–] | 0.02999 | 0.01707 | −0.03390 | −3.74 |
| $g_w$ [–] | −0.18200 | 0.00596 | −0.10000 | +13.77 |
| $T$ [°C] | 22.8617 | 0.0790 | 22.8793 | – |
| $u$ [–] | 0.5634 | 0.0016 | 0.5624 | – |
| $T_c$ [°C] | taken as 18 | 0 | 17.6937 | – |

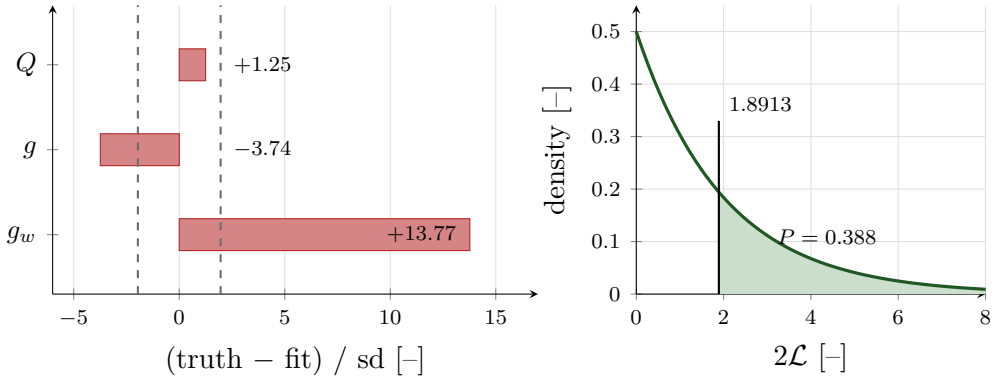



Figure 5.6: Two tests of one fit. Left: the entries' distance from the stated truth in the fit's own standard deviations, with dashed lines at ±1.96. Right: the misfit against its reference, the density $\frac{1}{2}e^{-x/2}$ of $\chi^2_2$, with the probability beyond the misfit shaded. The left test can be run only because the record is constructed, and the right one could not have seen what the left one shows.

The load is +1.25 sd away, the air path −3.74 sd and the coil **+13.77** sd. The hall air is met, 22.862°C against 22.879°C, and so is the valve. A fit can meet its readings and the temperature and be wrong about the coil by 14 of its own standard deviations, which is 7.33 of the sampling standard deviation 0.0112 measured in Section 5.6.5. The misfit test could not have seen this.

The rows say where to look. The load, the hall air and the valve are read directly, and they are the three that are met. The entries that are far away are the two whose rows run through $T_c$: the hall's difference $T - T_c$ and the coil's $T_c - T_w$. With the truth's own coil and valve, the coil's row gives 30.073 kW with the truth's supply air and 32.567 kW with $T_c = r$. Refitting this window with $T_c$ held at 17.7°C brings the true coil to +0.70 sd and the true air path to −0.25 sd, with the same readings. The modeling choice of step 0 is the only suspect left, and this graph cannot test it, because the one reading that might has a zero row here. Why the truth's

$T_c$ differs from $r$ is the question of Chapter 6. The misfit test is not a verdict on the truth: its degrees of freedom are the priors', in a drawing in which the choice is not up for question.

## 5.5 The report (step 6)

The sheet's forward line does not survive: the rated 0.938 is overturned, the fit puts the valve at a required 1.1246 at 17°C and 45 kW, and the record's valve sits on its stop there. The fit that says so meets its readings and the hall air, with a misfit at $P = 0.388$ that could not have said otherwise, and it is still 13.77 of its own standard deviations from the true coil, 7.33 of its sampling standard deviation. The answers below are reported under the sensor-true choice, which is the suspect.

### 5.5.1 The answers and their bands

Table 5.12 gives the answers in the order of Table 5.1. Each row is a question; the second column is the state route's answer with its spread, the third the draws beside it and the fourth the record's stated truth where it has one, a row for the check and not part of the answer.

Table 5.12: The report for the closed-loop hall's window, under the sensor-true choice. State route with sd; draws as median and sd.

| Question | State route | Draws | Truth |
|---|---|---|---|
| What does closing the loop change in the graph? | three entries left in both; $y_c$ a zero row, residual +0.0032 K; $g_w = -0.18200$ (0.00596) closed, −0.18277 (0.00657) open | – | $g_w = -0.10000$ |
| What is the plant now? | $G = 6.1827$ (0.1055), $K_Q = 0.1617$ (0.0028), $W = 13.3377$ (0.0794), $K_u = -7.0996$ (0.0199) | 6.1845 (0.1057), 0.1617 (0.0028), 13.3378 (0.0794), −7.0996 (0.0199) | +13.77 sd on $g_w$, −3.74 sd on $g$ |
| Lower the set point 1 K, closed | $T_c^* = 17$ (0), $T^* = 21.862$ (0.079), $u^* = 0.7512$ (0.0021) | 21.8605 (0.0790), 0.7512 (0.0021) | – |
| The same, open | closed posterior: $T_c^* = 17.000$ (0.0084), $T^* = 21.862$ (0.079), $u$ pinned 0.7512; own fit: $T_c^* = 17.0023$ (0.0119) | 17.0000 (0.0084), 21.8604 (0.0794) | – |
| A load step to 45 kW, closed | $T_c^* = 18$ (0), $T^* = 25.278$ (0.124), $u^* = 0.8435$ (0.0050) | 25.2763 (0.1243), 0.8435 (0.0050) | at 44.33 kW: $T = 25.330$, $u = 0.8338$ |
| The same, open | closed posterior: $T_c^* = 19.988$ (0.036), $T^* = 27.267$ (0.138), $u$ pinned 0.5634; own fit: $T_c^* = 19.9926$ (0.0394) | 19.9883 (0.0356), 27.2645 (0.1383) | – |
| Will the valve reach its stop at 17°C and 45 kW? | $u^* = 1.1246$ (0.0067); printed 1.0000, a normal tail at $\Phi(18.607)$ | 1.1246 (0.0067); printed 1.0000 | at 44.72 kW: on its stop, $u = 1.0000$ |

### 5.5.2 The entry in the commissioning log

The commissioning entry for the window records the loop as holding its reading, with a loop error of +0.003 K, and notes that this figure is no evidence about the sensor, since the loop consumes it. It strikes the sheet's line: 17°C at 45 kW is not available on this coil, whose required valve is above its stop, and the record's later step shows the loop losing its set point there. It records the coil as about 13.34 kW/K and the air path as about 6.18 kW/K, both with

spreads, and both flagged as resting on the discharge sensor being taken as true. A small misfit is not a clearance of that choice, since that test could not have seen it, and no gain or coil figure from this entry is to be used for tuning until the sensor question has been taken up.

## 5.6 The next measurement (step 7)

The report leaves two spreads at the node of the stop, the hall air and the valve at 17°C and 45 kW, and one suspect. A further reading is one more row, and what it would remove is known before its value is.

### 5.6.1 The preposterior of a candidate

A candidate $a$ is a row or a pair of rows with Jacobian $\mathbf{J}_a$. Its preposterior covariance is equation (4.73),

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1}, \qquad \mathrm{sd}_a(f) = \sqrt{\mathbf{g}_f^{\mathsf{T}}\boldsymbol{\Sigma}_a\,\mathbf{g}_f}, \tag{5.16}$$

with $\mathbf{g}_{T^*} = (0,\ -Q^*/G,\ 0)$ and $\mathbf{g}_{u^*} = (0,\ 0,\ -u^*)$ at the node. The spreads now are 0.1242 K and 0.0067. The gradients show what each spread is: that of $T^*$ is the air path's and that of $u^*$ is the coil's. The value of $T^*$ at this node is itself a what-if past the stop, and the row's spreads are compared here only as a measure of what is known of those two entries.

The candidates are the row of a reading on the closed graph: $\mathbf{g}_T^{\mathsf{T}}/\sigma_a$ for a sensor on the hall air, $\mathbf{g}_u^{\mathsf{T}}/\sigma_a$ for one on the valve, with the gradients (5.10), and $(1,\ 0,\ 0)/\sigma_a$ for one on the load. A reference discharge sensor is the row of $T_c$, whose gradient in equation (5.10) is zero. A further twenty minutes at 45 kW with the same sensors is three rows on one new entry: the window's own load, with its meter row, and $T$ and $u$ at that load. Equation (5.16) then runs with $\boldsymbol{\Sigma}^{-1}$ extended by a zero line for the new entry, which is marginalized out of the result. That window already exists in the record, steps 120 to 159, at a load of 44.45 kW.

### 5.6.2 The candidates

Table 5.13 ranks them: each row is a candidate, and the columns give each spread after it, by equation (5.16), with what it removes. The second hall sensor and the better load meter are permanent instruments; the reference sensors are one visit; the window at 45 kW is already in the record.

Table 5.13: Candidate readings for the spreads of $T^*$ and $u^*$ at 17°C, 45 kW, from 0.1242 K and 0.0067, in this graph and under the sensor-true choice.

| Candidate | $\mathrm{sd}_a(T^*)$ [K] | Removed [K] | $\mathrm{sd}_a(u^*)$ [–] | Removed [–] |
|---|---|---|---|---|
| Twenty minutes at 45 kW, the same sensors | 0.0691 | 0.0552 | 0.0037 | 0.0030 |
| Second hall sensor, twenty minutes (permanent) | 0.0919 | 0.0323 | 0.0067 | 0.0000 |
| Reference hall sensor 0.1 K, one reading | 0.1003 | 0.0239 | 0.0067 | 0.0000 |
| Load meter 0.2 kW, averaged (permanent) | 0.1185 | 0.0058 | 0.0034 | 0.0033 |
| Valve position sensor 0.002, averaged | 0.1242 | 0.0000 | 0.0059 | 0.0008 |
| Reference discharge sensor 0.05 K | 0.1242 | 0.0000 | 0.0067 | 0.0000 |

*In this graph the reference discharge sensor removes nothing*, because $T_c$ is pinned by the choice: its line in $\mathbf{J}_a$ is zero. The table is therefore only as good as that choice. It ranks readings for a plant in which the discharge sensor is true, and the check of Section 5.4.4 has made that the first thing to doubt.

### 5.6.3 The recommendation

Within this graph the ranking is plain. For the air path, twenty minutes at 45 kW removes the most, and it needs no new instrument and no visit, since that window already exists in the record. For the coil, the better load meter removes the most, 0.0033 against the window's 0.0030, because the 0.8821 of Section 5.2.2 makes the coil's spread mostly the load's, and it is the permanent instrument of the two. The recommendation is not to act on this ranking yet. The spreads it narrows are small against the coil's distance from the truth, so the entry defers the request, and Chapter 6 reranks the candidates with the sensor's choice opened.

### 5.6.4 What the analysis showed

The sheet read the trend as a clean bill and the rating as room on the valve. The analysis finds that the trend's flat error cannot speak about its own sensor, that the rating's headroom is not there, and that a fit which meets its readings and the hall air can still be far from the true coil, with a misfit test that could not have seen it. Closing the loop changed one thing in the count and two in the reading of it: the same three entries are left, but the what-if's band now sits on the valve rather than on the supply air, and the discharge reading that informed the coil in the sibling volume's open-loop reading has, at steady state and in window means, stopped informing anything. The coil is now read through the valve and the load, and through a supply air that the graph assumes. The loop buys the hall its protection from the disturbance, and it costs the fit its view of the one sensor the loop relies on.

### 5.6.5 Limits of the model

**Steady state.** The chapter says nothing about whether the loop gets to its operating point, how fast, or whether it stays stable as the coil and load change; Chapter 8 asks that.

**Static rows on window means.** The forty sensor errors are averaged as independent draws, and that $\sqrt{40}$ is fair; what fails is $T_c = r$ and static rows on window means. A study over 200 fresh records of the same plant and truth (seeds 1000 to 1199) measures it. The true window-mean $T_c$ moves with sd 0.0168 K across records, and the mean of the reading with only 0.0072 K against the 0.0158 of the sensor over $\sqrt{40}$: the loop holds the reading and passes the sensor's noise into the air, which $T_c = r$ does not allow. The load also wanders within a window that the rows read as steady. Together they move $\hat{g}_w$ with sd 0.0112 across records, against its own 0.00596. So the 13.77 is of the fit's own sd, and on the sampling sd the truth is 7.33 away. Across the records the distance in own sd is never below 8.87, so this record is not an unlucky one; and the refit at $T_c = 17.7$°C puts the truth at +0.70 sd.

**The sensor-true choice.** Every answer carries it. The bands here do not include it, and the check of Section 5.4.4 shows that it matters.

**The stop as a threshold.** The node of Section 5.3.5 asks whether the controller's row can hold, and it treats the stop as a line to cross. Past the line the plant is another regime, with the valve pinned and the set point lost, which Chapter 15 models as such.

**One coil conductance.** The coil is one conductance proportional to the valve's position. A valve whose characteristic is not linear, or a coil whose conductance changes with the flow, would move $g_w$ with the operating point, and the what-ifs at 0.8435 and above extrapolate.

**Constructed readings.** The record is constructed, generated from a stated truth, not field data. That is what makes the truth row possible, and the check of step 5 is one of the method on one record, one draw of the sensors' errors, with the study above as its only widening.

**The next chapter.** Chapter 6 frees the entry that this chapter has held at zero.

## 5.7 What to retain

**Definition 5.1** (Controller's row)**.** At steady state a controller with integral action holds a reading at its set point: $T_c + b = r$, with $b$ the sensor's bias. In the plant graph this is a row. The closed graph pins the set point, frees the actuator and adds this row; the open graph pins the actuator. The number of entries left is the same in both. A scenario node says which graph it belongs to, and at a stop the actuator is pinned again.

Three conclusions carry beyond the hall. Closing a loop moves a what-if's band from the controlled quantity to the actuator, and it keeps the disturbance from the controlled quantity at the actuator's expense. At steady state, in window means, a reading held at a set point has a zero row, and only another row on $T_c$ or $b$ opens it, so the loop's error is no evidence about the sensor. And a small misfit has been tested only inside the drawing it was given: where the readings are met by construction its degrees of freedom are the priors', and an assumption with no entry cannot fail it.

Table 5.14 lists the mistakes this chapter's hall invites. Each row is a reading of the record, what it costs here, and the correction. Table 5.15 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 5.14: Mistakes in reading a closed loop, and their corrections.

| Mistake | What it costs on this hall | Correction |
|---|---|---|
| A flat loop error read as a healthy sensor | $+0.0032$ K says nothing about the sensor | the reading has a zero row; the error is the loop's |
| The rating carried into a what-if | 0.938 against a required 1.1246 | fit the coil; its prior is one row among several |
| A zero spread read as knowledge | $T_c^*$ with sd 0 closed, $u^*$ with sd 0 open | it is the pin or the sensor-true choice, not a measurement |
| Agreement of open and closed fits read as confirmation | $-0.18200$ and $-0.18277$ | both take the sensor as true |
| A small misfit read as a verdict on the truth | $P = 0.388$ with $g_w$ at 13.77 sd | its degrees of freedom are the priors'; it could not have seen this |
| A printed 1.0000 read as certainty | $\Phi(18.607)$ | a normal tail under a choice outside the band |

Table 5.15: The terms of the closed-loop hall and where each is defined.

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| PI controller | eq. (5.1) | – | the sampled loop | 5.1.1 |
| Controller's row | $T_c + b = r$ | °C | the reading held at the set point | 5.1.1 |
| Physical rows | eq. (5.4) | K, – | hall and coil balances | 5.1.2 |
| Process gain | $K_u = -(T_c - T_w)/u$ | K | the valve on the supply air | 5.3.2 |
| Sensor-true choice | $b = 0$, so $T_c = r$ closed | – | three entries; the suspect | 5.1.2 |
| Reduced state | $(Q, g, g_w)$ | kW, – | the same in both graphs | 5.1.4 |
| Closed map | eq. (5.6) | °C, – | $r$ pinned, $u$ follows | 5.1.4 |
| Open map | eq. (5.7) | °C | $u$ pinned, $T_c$ follows | 5.1.4 |
| Zero row | $\partial T_c/\partial \mathbf{z} = \mathbf{0}$, eq. (5.10) | – | the loop consumes its reading | 5.2.1 |
| State route | $f(\hat{\mathbf{z}})$, $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$ | varies | the reported value | 5.3 |
| Reweighted draws | weights to the exact posterior | varies | the other route | 5.3 |
| Node gradients | eq. (5.14) | varies | where the band lands | 5.3.3 |
| Stop tail | eq. (5.15) | – | the row cannot hold | 5.3.5 |
| Misfit | $2\mathcal{L}(\hat{\mathbf{z}})$ against $\chi^2_2$ | – | the priors against the fit | 5.4.4 |
| Preposterior | eq. (5.16) | varies | spread after a candidate | 5.6.1 |

## 5.8 Exercises

**Exercise 5.1.** *Formulation.* A tank has level $h$, an inflow $q$ it does not set and an outflow $k\,u\,\sqrt{h}$ through a valve at position $u$, with a level sensor read by a PI controller holding set point $r$. Write the graph in both forms and count what is left. (Entries $(h, u, q, k)$ and one physical row, $q = k\,u\,\sqrt{h}$. Open: $u$ pinned, one row on three free entries, two left. Closed: $r$ pinned, $u$ free, and the controller's row $h + b = r$, two rows on four free entries, two left, $(q, k)$ in both. Closed with $b = 0$, the level reading has a zero row, and the valve's reading informs $k$.)

**Exercise 5.2.** *Derivation.* Show why the closed graph's discharge reading has a zero row, and what opens it. (Its row is $(T_c + b - y_c)/\sigma$, and on the set the controller's row gives $T_c + b = r$, so the row is the constant $(r - y_c)/\sigma$ whatever the state. Its gradient on $(Q, g, g_w)$ is zero by equation (5.10), its outer product in $\mathbf{J}^\mathsf{T}\mathbf{J}$ is zero and it adds information 0. Freeing $b$ does not open that row either; only another row on $T_c$ or $b$ does. Its residual is the loop's error, +0.0032 K.)

**Exercise 5.3.** *Derivation.* Derive the gradient of $u^*$ at a closed node, with the load free and with it pinned, and read its spread at the stop. (From $u^* = Q\,\mathrm{e}^{-g_w}/(W_{\text{rated}}(r^* - T_w))$, $\partial u^*/\partial Q = u^*/Q$ and $\partial u^*/\partial g_w = -u^*$, with no term in $g$; pinning the load removes the first. At 17°C and the present load, $\mathbf{g}_{u^*} = (0.02499,\ 0,\ -0.75122)$, sd 0.0021. At 17°C and 45 kW, $\mathbf{g}_{u^*} = (0,\ 0,\ -1.1246)$, and the spread is $1.1246 \times 0.00596 = 0.0067$.)

**Exercise 5.4.** *Derivation.* The open $T_c^*$ at the set-point move, on the closed $\boldsymbol{\Sigma}$, has a smaller spread than either of its diagonal terms would give. Read the reason from $\mathbf{g}$ and $\boldsymbol{\Sigma}$. (Its gradient is $(0.09981,\ 0,\ -3)$, opposite signs on $Q$ and $g_w$, and their correlation is +0.8821, so the cross term is negative. The budget is $0.00025 + 0.00032 - 0.00050\,\mathrm{K}^2$, and the spread is 0.0084 K.)

**Exercise 5.5.** *Application.* At 40 kW, where does the valve stand at $r^* = 18$°C, and at what set point does it first reach its stop? Set that beside the rated coil. ($u^* = 40/(W(18 - 14))$, at the fit's mean 0.7498, sd 0.0045. The valve reaches $u^* = 1$ at $r^* = T_w + Q^*/W$, which at the fit's mean is 16.9990°C, under the sensor-true choice; at the rated coil it is 16.5000°C.)

# Chapter 6

# Protected Quantities and Chance Constraints

## 6.1 A limit the loop's error cannot see

The commissioning engineer has the loop's trend open and a request on the desk: 45 kW in the hall, and then 60 kW, with the hall air below its limit of 27.5°C. The trend gives no cause for concern. Over the twenty minutes of the record at about 30 kW the discharge reading averages 18.0032°C against a set point of 18°C, an error of +0.0032 K, and Chapter 5's fit of those minutes passed its misfit test with $P = 0.388$. That fit puts the hall at 25.278°C at 45 kW with a standard deviation of 0.124 K. The engineer would conclude that the sensor and the loop are sound, and that the spread on the temperature is what stands between the set point and the limit.

The same fit missed the coil by 13.77 of its own standard deviations, and it took the discharge sensor as true to get there. A controller holds its reading at the set point, not the air the reading is meant to measure, so the trend would look the same whatever the sensor's offset. What the trend and the misfit test leave open is how far the set point may be trusted to stand for the air, and a loop has a second limit besides: the valve, whose stop ends the loop's hold on anything. The question of this chapter is

*what must be protected, with what probability, and what a sensor the loop cannot see does to it.*

**Jargon: protected quantity, chance constraint, stated risk, joint chance constraint, back-off, window of set points, sensor offset, calibration tolerance, reference sensor, observability, importance weights, effective sample size**

A *protected quantity* is a quantity of the plant that must stay inside a limit for the plant to do its job. A *chance constraint* requires that limit to hold with a stated probability given the evidence; the *stated risk* $\alpha$ is one minus that probability. A *joint chance constraint* requires several limits to hold together, in one event, with one probability. The *back-off* is the distance a quantity's mean must stand inside its limit because the quantity is uncertain. The *window of set points* is the range of set points that meets a joint chance constraint at one load; it is not the stretch of the record that the fit reads, which this chapter calls the record's twenty minutes. A *sensor offset* (in the framework, a bias entry of the state) is a steady error of a sensor, and its *calibration tolerance* is the spread the sensor's calibration allows. A *reference sensor* is a better instrument brought for a short check of another. *Observability* is whether the readings determine an entry of the state, read from the columns of the Jacobian and from the entry's posterior spread against its prior spread. *Importance weights* turn draws from an approximate distribution into draws from the exact one, and the *effective sample size* is the number of equally weighted draws they are worth.

**Key idea**

A chance constraint is a threshold question on a function of the state at a scenario node. For a normal posterior it becomes the mean plus a back-off, and the back-off is a quantile times $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$. A sensor offset under feedback is an entry of the state: it enters the protected temperature with gradient $-1$, and the data cancel most of it only near the condition they were taken at. A loop has two protected quantities, so a load has a window of set points, and the load at which the window closes is the answer to a request.

### 6.1.1 Protected quantities and the general chance constraint

Chapter 2 wrote a requirement as the threshold form (2.12), and Chapter 4 read it at one scenario node with an input entering additively. The form a loop needs has a limit above the quantity, a pinned input $r$ and a stated risk $\alpha$. It is equation (2.12) with its $f$ taken as $-f(\mathbf{z};r)$, its $f^*$ as $-f^*$ and $\kappa = 1-\alpha$:

$$\Pr\big(f(\mathbf{z};r) \le f^* \mid \mathbf{y}\big) = \int \mathbf{1}\big[f(\mathbf{z};r) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z} \;\ge\; 1-\alpha. \tag{6.1}$$

This is the general chance constraint, defined here once. The quantity $f$ may be an entry of the state or any function of it, the input $r$ is pinned at a scenario node and is the thing the engineer may choose, and the limit and the risk are the plant's policy. Several limits held together are the same integral with the indicator of an intersection.

For a normal posterior the integral has a closed form. Linearizing $f$ at the MAP state with the gradient $\mathbf{g}(r) = \partial f/\partial \mathbf{z}$ at $\hat{\mathbf{z}}$, the delta method (4.53) gives $f \mid \mathbf{y} \approx \mathcal{N}\big(f(\hat{\mathbf{z}};r),\ \mathbf{g}(r)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(r)\big)$. The probability in equation (6.1) is then a normal tail,

$$\Pr\big(f(\mathbf{z};r) \le f^* \mid \mathbf{y}\big) = \Phi\left(\frac{f^* - f(\hat{\mathbf{z}};r)}{\sqrt{\mathbf{g}(r)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(r)}}\right), \tag{6.2}$$

and since $\Phi$ is increasing, applying $\Phi^{-1}$ to both sides of the constraint and multiplying through by the spread gives

$$f(\hat{\mathbf{z}};r) + \Phi^{-1}(1-\alpha)\,\sqrt{\mathbf{g}(r)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(r)} \;\le\; f^*. \tag{6.3}$$

The second term on the left is the back-off that Section 4.3.8 named for one node. Two conditions come with its general form. The first is the normality of $f$: the state may be normal while a nonlinear $f$ of it is not, and then equation (6.3) is an approximation whose error sits in the tail it is read from. The second is the dependence of the gradient on $r$. Where $r$ enters additively the gradient does not contain it, the spread is one number, and the constraint inverts in closed form, which is equation (4.66). Where $r$ does not enter additively the gradient, and so the back-off, changes with $r$, and equation (6.3) has to be solved for $r$ with the spread recomputed at every trial value. This chapter meets both cases on one plant. Choosing one set point under such a constraint is the one-move case of the problem of Chapter 7.

### 6.1.2 The closed-loop hall with its offset freed (step 0)

The plant is Chapter 5's at its steady operating point: the hall air $T$, the supply air $T_c$ leaving the coil, the valve $u$, the load $Q$, the path conductance $G = G_{\mathrm{ref}}\mathrm{e}^{g}$ with $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$, and the coil's conductance $W = W_{\mathrm{rated}}\mathrm{e}^{g_w}$ with $W_{\mathrm{rated}} = 16\,\mathrm{kW/K}$ fully open, on chilled water pinned at $T_w = 14°\mathrm{C}$. The set point $r$ is pinned and the valve is free, as in any closed-loop graph. One modeling choice of that chapter is released here.

**The choice released.** Chapter 5 took the discharge sensor as true, which made the supply air the set point. The integrator of the PI controller forces the reading, not the air, to the set point at steady state. With a sensor offset $b$, so that the reading of the air is $T_c + b$, the controller's row is $T_c + b = r$, and the offset passes into the true supply temperature with the opposite sign, $T_c = r - b$. This is what lets the graph carry an error the loop cannot report, and it adds an entry. The offset is taken as one constant for the record and for any set point that follows; a sensor whose error drifts or depends on the temperature would break this (Section 6.6.5).

**The physical rows.** The full state has seven entries, $(T,\ T_c,\ u,\ Q,\ g,\ g_w,\ b)^{\mathsf{T}}$. Written with the tolerances that the limit $\sigma_p \to 0$ removes, the three rows are

$$\begin{aligned} r_{\mathrm{h}} &= \left[T - T_c - Q/(G_{\mathrm{ref}}\mathrm{e}^{g})\right]/\sigma_p && \text{hall balance,} \\ r_{\mathrm{c}} &= \left[u - Q/(W_{\mathrm{rated}}\mathrm{e}^{g_w}(T_c - T_w))\right]/(0.01\,\sigma_p) && \text{coil balance,} \\ r_{\mathrm{l}} &= \left[T_c + b - r\right]/\sigma_p && \text{controller's row at steady state,} \end{aligned} \tag{6.4}$$

the valve's row carrying a smaller tolerance because the valve is dimensionless. On the reduced state of Section 4.2.8 the free entries are

$$\mathbf{z} = (Q,\ g,\ g_w,\ b)^{\mathsf{T}}, \tag{6.5}$$

in kW, two dimensionless entries and K, Chapter 5's three with the offset added, and the dependent quantities follow from the rows,

$$T_c = r - b, \qquad T = r - b + \frac{Q}{G_{\mathrm{ref}}\mathrm{e}^{g}}, \qquad u = \frac{Q}{W_{\mathrm{rated}}\mathrm{e}^{g_w}\,(r - b - T_w)}. \tag{6.6}$$

**The graph.** Figure 6.1 draws these rows. The chain along the middle runs from the pinned chilled water on the left through the coil and valve (green, carrying $g_w$ and the free valve) to the supply air (blue, fixed by the rows), the hall air (blue, its path carrying $g$) and the free load on the right. Above the chain the discharge sensor adds the free offset $b$ (green) to the supply air, and the controller's row compares the result with the pinned set point and moves the valve. The dashed arrows are the three readings that remain rows. The figure's point is the top line: the loop's own reading passes through the sensor into the controller's row and stops there. Its comparison with $r$ is the loop's error, which holds no entry of the state.

**The scenario node.** Every what-if of this chapter is closed loop. It pins the set point at $r$ and the load at $Q^*$, leaves the valve free and determined by the rows, and shares $g$, $g_w$ and $b$ with the record:

$$T^*(r) = r - b + \frac{Q^*}{G_{\mathrm{ref}}\mathrm{e}^{g}}, \qquad u^*(r) = \frac{Q^*}{W_{\mathrm{rated}}\mathrm{e}^{g_w}\,(r - b - T_w)}. \tag{6.7}$$

This form holds only while $u^*(r) \le 1$. A value above 1 is the demand the valve cannot meet: at the stop the actuator is pinned again and the set point is lost, which is why the stop is a limit. This chapter calls the node *void* there: equation (6.7) then describes no plant. Chapter 15 takes up what the plant does there.

### 6.1.3 The control engineer's questions (step 1)

The engineer asks six questions of the record, each written over the reduced state with the evidence on the right of the bar. Table 6.1 lists them. The first column is the question in the engineer's words, the second the general form of its category, the third its mathematical form on this hall, the fourth its category, the fifth the computation that answers it and the last the subsection. The first asks for moments of the state; the other five are the chance constraint, single or joint, read one way or inverted. The limit is $T_{\max} = 27.5°\mathrm{C}$ and the stated risk, unless a row says otherwise, is $\alpha = 0.05$.

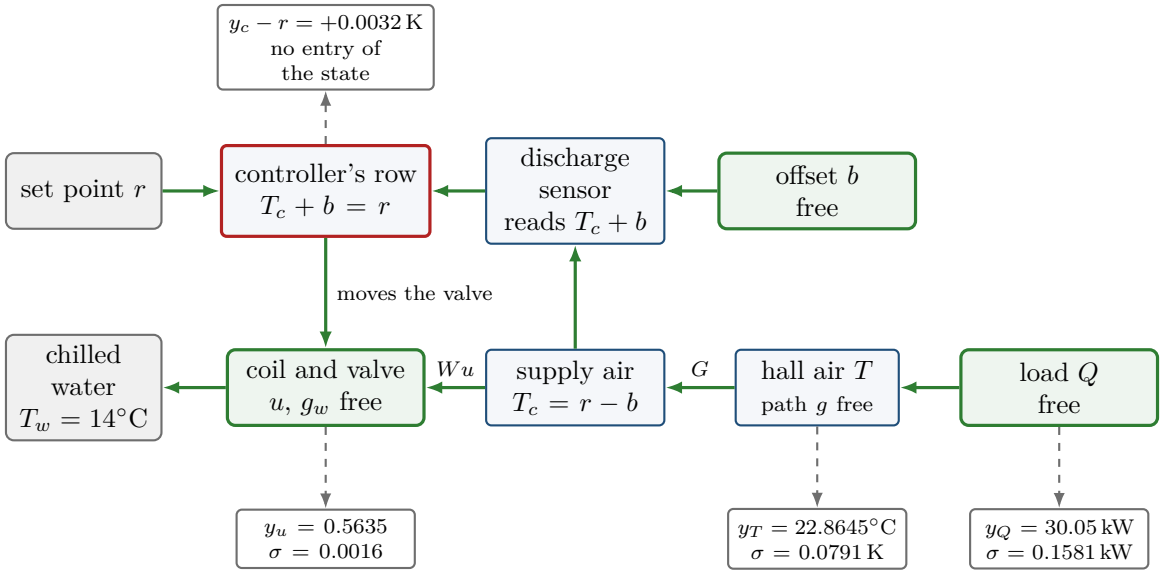



Figure 6.1: The closed-loop hall with the discharge sensor's offset freed. Free (green): the load $Q$, the coil and valve with $g_w$, and the offset $b$; the path's $g$ is a free entry carried on the hall air's node. Blue: quantities fixed by the rows (6.4). Pinned (gray): the chilled water and the set point. The controller's row (red) consumes the discharge reading; what is left of it is the loop's error, which holds no entry. Observed (dashed): the three reading rows. Priors on $g$, $g_w$ and $b$ are rows not drawn.

Table 6.1: The questions asked of the closed-loop hall with the offset freed. $\mathbf{y}$ is the three readings and the three priors; $T^*$ and $u^*$ follow equation (6.7).

| Control engineer's question | General form | Mathematical form | Category | Computation | Answered in |
|---|---|---|---|---|---|
| What does an offset the loop cannot see do to the plant? | moments, (2.10) | $p(b, g, g_w \mid \mathbf{y})$, $p(T^*, u^* \mid \mathbf{y})$ | estimation | posterior | 6.3.1 |
| What must stay inside its limit, with what probability? | chance constraint, (6.1) | $\Pr(T^*(r) \le 27.5°\mathrm{C} \mid \mathbf{y}) \ge 1 - \alpha$ | prediction under a threshold | posterior | 6.3.2 |
| What is the highest set point that holds it? | the same, inverted | $r_{95}$ | prediction under a threshold | posterior | 6.3.3 |
| Is the valve a protected quantity too? | the same | $\Pr(u^*(r) \le 1 \mid \mathbf{y}) \ge 1 - \alpha$ | prediction under a threshold | posterior | 6.3.4 |
| Which set points hold both? | joint chance constraint | $\Pr(T^* \le 27.5°\mathrm{C},\ u^* \le 1 \mid \mathbf{y}) \ge 1 - \alpha$ | prediction under a threshold | posterior | 6.3.5 |
| At what load does no set point hold? | the same, over the load | the window's width against $Q^*$ | prediction under a threshold | posterior | 6.3.5 |

### 6.1.4 The evidence as rows (step 2)

The evidence is Chapter 5's: the means of steps 40 to 79 of the record, twenty minutes at the set point of 18°C, each reading's standard deviation divided by the square root of the forty steps. They are $y_T = 22.8645$°C with $\sigma = 0.0791$ K, $y_u = 0.5635$ with $\sigma = 0.0016$, and $y_Q = 30.05$ kW with $\sigma = 0.1581$ kW. The priors on $g$ and $g_w$ are that chapter's, the handbook's $g \sim \mathcal{N}(0.22314,\ 0.30^2)$ and the rated coil's $g_w \sim \mathcal{N}(0,\ 0.15^2)$; the load has none.

**The modeling choice at the new row.** The offset has the prior $b \sim \mathcal{N}(0,\ 0.20^2)$ in K, the sensor's calibration tolerance. This is a stated choice. It buys a state the fit can solve and a band that carries the offset; it fails if the sensor is further out of calibration than its tolerance allows, and Section 6.4.3 shows that the data do not correct it. Each reading is taken as an independent normal row given the state, as in Chapter 5.

**The reading the loop consumes.** The discharge reading is not among the rows. Its model is $y_c = T_c + b$, and on the reduced state the controller's row makes $T_c + b$ equal to $r$, whatever $b$ is. The row would read $y_c - r$, with a zero line in every column, so the loop's error of $+0.0032$ K is the same under any offset, and its information about the offset is zero. The trend of the opening scene cannot see $b$ for this reason, and for no reason to do with the size of $b$.

**Which row acts on which entry.** By equation (6.6) the hall reading depends on $Q$, $g$ and $b$, the valve reading on $Q$, $g_w$ and $b$, and the meter on $Q$. With $*$ for an entry that can be nonzero, the $6 \times 4$ Jacobian has the pattern

$$\mathbf{J} = \begin{bmatrix} * & * & 0 & * \\ * & 0 & * & * \\ * & 0 & 0 & 0 \\ 0 & * & 0 & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} \text{hall reading } y_T \\ \text{valve reading } y_u \\ \text{load meter } y_Q \\ \text{prior on } g \\ \text{prior on } g_w \\ \text{prior on } b \end{matrix} \tag{6.8}$$

with columns $Q, g, g_w, b$. Two facts are read here before any number. The column of $b$ has two reading rows, and each of those rows also holds a column of its own, $g$ for the hall reading and $g_w$ for the valve reading: whatever $b$ does to a reading, a change of that entry can undo. The three reading rows act on four entries, so their rank is at most three, and only the prior rows separate the offset from the two conductances.

## 6.2 The posterior of the state (step 3)

All six questions read one posterior, that of the reduced state (6.5) under the six rows of step 2. It is Chapter 5's fit with one more column and one more row.

### 6.2.1 The rows, the normal equations and the covariance

Stacked in the order of the pattern (6.8), each divided by its standard deviation, the rows on the reduced state form the residual vector and the objective

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T(\mathbf{z}) - y_T)/0.0791 \\ (u(\mathbf{z}) - y_u)/0.0016 \\ (Q - y_Q)/0.1581 \\ (g - 0.22314)/0.30 \\ g_w/0.15 \\ b/0.20 \end{bmatrix}, \qquad \mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \tag{6.9}$$

with $T(\mathbf{z})$ and $u(\mathbf{z})$ from equation (6.6) at $r = 18°\mathrm{C}$. The Jacobian is the reduced one of equation (4.47). The hall reading's line carries $\partial T/\partial \mathbf{z} = (1/G,\ -Q/G,\ 0,\ -1)$ and the valve reading's carries $\partial u/\partial \mathbf{z} = (u/Q,\ 0,\ -u,\ u/(r - b - T_w))$, each over its sensor's standard deviation. Each iteration solves

$$(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \qquad \mathbf{z}^{(k+1)} = \mathbf{z}^{(k)} + \Delta\mathbf{z}, \tag{6.10}$$

the Gauss–Newton step on these normalized rows, whose information matrix is $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$; with unnormalized rows it is $\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$. The rows are nonlinear through $\mathrm{e}^{g}$, $\mathrm{e}^{g_w}$ and $1/(r - b - T_w)$, so the curvature term of equation (4.44) is kept in the covariance,

$$\mathbf{\Sigma} = \left(\mathbf{J}^{\mathsf{T}}\mathbf{J} + \mathbf{C}\right)^{-1}\big|_{\hat{\mathbf{z}}}, \tag{6.11}$$

and $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$. The Gauss–Newton form without $\mathbf{C}$ is set beside it in Section 6.4.1.

### 6.2.2 The fit of the record

**The iterations.** Table 6.2 runs the normal equations (6.10) from the meter's load, the handbook's $g$, the rated coil and no offset. Each row is an iteration; the columns give the four entries, the objective and the norm of its gradient. The objective falls from 1823 to below one in two steps, and the valve reading, whose standard deviation is the smallest, does most of that.

Table 6.2: Gauss–Newton iterations of the closed-loop hall with the offset free, equation (6.10), on the reduced state.

| $k$ [–] | $Q$ [kW] | $g$ [–] | $g_w$ [–] | $b$ [K] | $\mathcal{L}$ [–] | $\lVert\mathbf{J}^{\mathsf{T}}\mathbf{r}\rVert$ [–] |
|---|---|---|---|---|---|---|
| 0 | 30.0500 | 0.22314 | 0.00000 | 0.00000 | 1823.0166 | $1.82 \times 10^{4}$ |
| 1 | 30.0590 | −0.00395 | −0.18534 | 0.05718 | 22.2560 | $2.39 \times 10^{3}$ |
| 2 | 30.0581 | 0.01799 | −0.16770 | 0.05749 | 0.9022 | $2.06 \times 10^{1}$ |
| 3 | 30.0581 | 0.01833 | −0.16764 | 0.05709 | 0.9005 | $3.66 \times 10^{-3}$ |
| 4 | 30.0581 | 0.01833 | −0.16764 | 0.05708 | 0.9005 | $8.67 \times 10^{-6}$ |
| 5 | 30.0581 | 0.01833 | −0.16764 | 0.05708 | 0.9005 | $1.67 \times 10^{-7}$ |

**The Jacobian at the MAP state.** In the columns $Q, g, g_w, b$ and the rows of equation (6.9),

$$\mathbf{J}(\hat{\mathbf{z}}) = \begin{bmatrix} 2.06989 & -62.21703 & 0 & -12.64911 \\ 11.85492 & 0 & -356.33616 & 90.37367 \\ 6.32456 & 0 & 0 & 0 \\ 0 & 3.33333 & 0 & 0 \\ 0 & 0 & 6.66667 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}, \tag{6.12}$$

the first column in $\mathrm{kW}^{-1}$, the last in $\mathrm{K}^{-1}$, the others dimensionless. The zeros are the pattern's. The first two rows name the directions the readings cannot see. On the hall reading's line, a move along $(\Delta g, \Delta b) \propto (-12.64911,\ 62.21703)$ changes nothing; on the valve reading's line, a move along $(\Delta g_w, \Delta b) \propto (90.37367,\ 356.33616)$ changes nothing. Moving the offset along both leaves the two readings where they are, and only the prior rows object.

**The covariance.** The Laplace covariance (6.11) and the correlations are, in the same order,

$$\mathbf{\Sigma} = \begin{bmatrix} 0.02496 & 0.00088 & 0.00076 & -0.00027 \\ 0.00088 & 0.00177 & -0.00183 & -0.00732 \\ 0.00076 & -0.00183 & 0.00235 & 0.00913 \\ -0.00027 & -0.00732 & 0.00913 & 0.03605 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0.13272 & 0.09949 & -0.00899 \\ 0.13272 & 1 & -0.89555 & -0.916 \\ 0.09949 & -0.89555 & 1 & 0.99242 \\ -0.00899 & -0.916 & 0.99242 & 1 \end{bmatrix},$$

in $\mathrm{kW}^2$, $\mathrm{K}^2$ and mixed units. The two large correlations read by sign. A larger offset means colder true supply air. The valve reading must still be met with less temperature difference across the coil, so the coil must be better: $\mathrm{corr}(g_w, b) = +0.992$. The hall reading must still be met with a larger rise from colder air, so the path must be worse: $\mathrm{corr}(g, b) = -0.916$. Said the other way, a warmer true supply is taken up by a conductance.

**The MAP state.** Table 6.3 gives the fit beside Chapter 5's. Each row is an entry; the columns give that chapter's value and spread with the offset fixed at zero, the value and spread with it free, the prior's spread and each fit's distance from the stated truth of the constructed record, in its own standard deviations. The offset's value in that truth is left to Section 6.4.4.

Table 6.3: The MAP state with the offset fixed at zero (Chapter 5) and free, with the posterior spreads $\sqrt{\Sigma_{ii}}$. The last row is the true supply air, $T_c = r - b$, a function of the last entry.

| | Offset fixed at 0 | | | Offset free | | | |
|---|---|---|---|---|---|---|---|
| Entry [unit] | $\hat{z}_i$ | $\sqrt{\Sigma_{ii}}$ | to truth [sd] | $\hat{z}_i$ | $\sqrt{\Sigma_{ii}}$ | to truth [sd] | Prior sd |
| $Q$ [kW] | 30.0585 | 0.1580 | +1.25 | 30.0581 | 0.1580 | +1.25 | none |
| $g$ [–] | 0.02999 | 0.01707 | −3.74 | 0.01833 | 0.04210 | −1.24 | 0.30 |
| $g_w$ [–] | −0.18200 | 0.00596 | +13.77 | −0.16764 | 0.04846 | +1.40 | 0.15 |
| $b$ [K] | 0 | fixed | – | 0.0571 | 0.1899 | +1.28 | 0.200 |
| $T_c$ [°C] | 18 | fixed | – | 17.9429 | 0.1899 | −1.31 | – |

Freeing one entry changes the table in three ways. The offset's posterior spread is 0.1899 K against a prior of 0.200 K, which leaves 90.1 % of the prior's variance. The readings alone are not what removes the rest. With $g$ and $g_w$ made flat, 100.000 % is kept and $\hat{b} = 0$, so the three readings carry nothing on the offset; what is removed comes through the priors on the coil and the path. It amounts to $2.74\,\mathrm{K}^{-2}$ against the prior's $25\,\mathrm{K}^{-2}$, as much as a stand-alone reading of 0.60 K, and the percentage is itself prior-dependent: 31.9 % is kept at a prior of 0.80 K. In the framework's term the offset is unobservable from these readings. What moves it from zero to 0.0571 K is the prior rows on $g$ and $g_w$ pulling along the ridge. The spreads on $g$ and $g_w$ grow 2.5- and 8-fold, since they now share the offset's freedom. And the truth now lies inside the fit: the coil's distance of 13.77 falls to 1.40.

**The misfit.** The rows at the MAP state are −0.037, −0.021 and 0.051 for the three readings, and −0.683, −1.118 and 0.285 for the priors on $g$, $g_w$ and $b$. Then $2\mathcal{L}(\hat{\mathbf{z}}) = 1.8010$ on $6 - 4 = 2$ degrees of freedom, with $\Pr(\chi^2_2 \geq 1.8010) = 0.406$, against 0.388 for Chapter 5's fit on $5 - 3 = 2$. The misfit test cannot tell the two fits apart, although one of them is 13.77 standard deviations from the truth on the coil: each new row brought its new entry, and the readings are met in both.

## 6.3 The answers (step 4)

Every answer is read from the pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 6.2.2 by the state route: the node's mean and gradient, the quadratic form $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$, and a normal tail or, for the joint event, a bivariate normal integral done by deterministic quadrature. That value is the one reported. The route with reweighted draws is set beside the probabilities in Table 6.7 and tested in Section 6.4.2.

### 6.3.1 What the offset does to the plant

**Control engineer's question.** What does an offset the loop cannot see do to the plant?

**General form.** An estimation question reads moments, equation (2.10), of an entry or of a derived quantity $f(\mathbf{z})$ under the posterior, with the variance $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ of the delta method (4.53).

**Specialized form.** Equation (2.10) with $\mathbf{z}$ the reduced state (6.5) and $f$ one of the offset entry $b$, or $T^*$ or $u^*$ of equation (6.7). The gradients at the node are

$$\mathbf{g}_{T^*} = \begin{bmatrix} 0 \\ -Q^*/G \\ 0 \\ -1 \end{bmatrix}, \qquad \mathbf{g}_{u^*} = \begin{bmatrix} 0 \\ 0 \\ -u^* \\ u^*/(r-b-T_w) \end{bmatrix}, \tag{6.13}$$

with a zero in the column of $Q$ because the node pins the load. The offset enters the hall air with gradient $-1$: it passes in full, and that is the whole of what feedback does with it. The band is then not simply wider, because the hall reading's line of equation (6.12) has the same shape as $\mathbf{g}_{T^*}$. With $\lambda = Q^*/\hat{Q}$ the ratio of the node's load to the measured one,

$$\mathbf{g}_{T^*} = \lambda \frac{\partial T}{\partial \mathbf{z}} - \frac{\lambda}{G}\,\mathbf{e}_Q + (\lambda - 1)\,\mathbf{e}_b, \tag{6.14}$$

with $\mathbf{e}_Q$ and $\mathbf{e}_b$ the unit vectors of those columns. The first term is the combination the hall reading fixes and the second the one the meter fixes. Only the third is left to the offset's own spread, and its weight is $\lambda - 1$: zero at the measured load, about one half at 45 kW and about one at 60 kW. At the set point of the record the valve's gradient is, in the same way, a multiple of the valve reading's line in the columns of $g_w$ and $b$, so at $r = 18\,^{\circ}\mathrm{C}$ the offset leaves the valve's spread alone.

**Building blocks.** The fit and $\boldsymbol{\Sigma}$ of Section 6.2.2; the gradients (6.13); the budget of equation (4.54), one term per entry and the cross terms.

**The plant's numbers.** At $r = 18\,^{\circ}\mathrm{C}$ and $Q^* = 45$ kW the gradients are

$$\mathbf{g}_{T^*} = (0,\ -7.36377,\ 0,\ -1)^{\mathsf{T}}, \qquad \mathbf{g}_{u^*} = (0,\ 0,\ -0.84349,\ 0.21393)^{\mathsf{T}}.$$

Table 6.4 sets the two fits beside each other at that node. Each row is a quantity; the columns give its value with its spread with the offset fixed at zero and free, the ratio of the spreads, and the stated truth. The result goes against what the 13.77 led one to expect. At this prior, freeing the offset rescues the parameters and barely moves the operating answers: the hall air moves by a few hundredths of a kelvin and its spread grows by a quarter, while the valve does not change in the digits printed.

Table 6.4: The node at $r = 18\,^{\circ}\mathrm{C}$, $Q^* = 45$ kW, closed loop, in the two fits. Standard deviations in parentheses.

| Quantity [unit] | Offset fixed at 0 | Offset free | Ratio of spreads [–] | Truth |
|---|---|---|---|---|
| $T^*$ [°C] | 25.2784 (0.1242) | 25.3067 (0.1559) | 1.26 | 25.4586 |
| $u^*$ [–] | 0.8435 (0.0050) | 0.8435 (0.0050) | 1.00 | 0.8401 |

The budget of $\mathrm{Var}(T^*)$ shows the cancellation directly. In $\mathrm{K}^2$ the column of $g$ gives 0.09609, the column of $b$ gives 0.03605, and the cross terms give $-0.10782$, for a total of 0.02432. The

cross term is negative because both gradient entries are negative and their covariance $\Sigma_{gb}$ is negative as well: a large offset and a poor path come together, and they push the hall air in opposite directions. The two diagonal terms taken alone would overstate the variance several times.

Stated plainly, the offset passes into the true temperature in full, yet near the measured condition the data cancel most of it, and the cancellation weakens with distance from that condition, as $\lambda - 1$ in equation (6.14). The spread of $T^*$ with the offset free against fixed is 0.0829 against 0.0828 K at 30 kW, 0.1559 against 0.1242 K at 45 kW and 0.2512 against 0.1656 K at 60 kW.

The cancellation is structural. From equation (6.7), at the record's set point and with $\lambda = Q^*/Q$, there is an exact identity,

$$T^* = \lambda\, T - (\lambda - 1)\, r + (\lambda - 1)\, b, \tag{6.15}$$

in which the reading fixes $T$ and $r$ is pinned, so the spread left over is $|\lambda - 1|$ times the spread of $b$, 0.0944 K at 45 kW. The difference of the two fits' variances there is $0.00889\,\mathrm{K}^2$ against $(\lambda - 1)^2\Sigma_{bb} = 0.00891\,\mathrm{K}^2$. What is small about it is conditional: the prior is 0.2 K and $\hat{b} = 0.057$ K. At a prior of 0.80 K, $r_{95}$ stands 0.39 K below the sensor-true fit's at 45 kW and 0.86 K below at 60 kW, the second figure a void reading, since both set points lie below that load's floor. The band must carry the offset; at this prior the numbers show it costs less than the coil's 13.77 suggested.

### 6.3.2 The chance constraint on the hall air

**Control engineer's question.** What must stay inside its limit, with what probability?

**General form.** The chance constraint (6.1), the threshold form (2.12) for a limit above the quantity, at a scenario node with the input $r$ pinned.

**Specialized form.** Equation (6.1), itself equation (2.12), with $f(\mathbf{z}; r) = T^*(r)$ of equation (6.7), $f^* = 27.5$°C and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal of step 3. By equation (6.13) the set point is not in the gradient, since $r$ enters $T^*$ additively, and this is Section 4.3.8's result with one more column. The tail is equation (6.2) with one spread at every $r$,

$$\Pr(T^*(r) \le 27.5\text{°C} \mid \mathbf{y}) = \Phi\left(\frac{27.5 - T^*(r; \hat{\mathbf{z}})}{\sqrt{\mathbf{g}_{T^*}^{\mathsf{T}}\, \boldsymbol{\Sigma}\, \mathbf{g}_{T^*}}}\right). \tag{6.16}$$

The normality of $T^*$ is the choice here: $T^*$ is curved in $g$ through $\mathrm{e}^{-g}$, and the draws of Section 6.4.2 test what that does in the tail.

**Building blocks.** $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$, the gradient of equation (6.13) and $\Phi$.

**The plant's numbers.** At the present set point of 18°C the hall air at 45 kW has mean 25.307°C and spread 0.156 K, and the probability is >0.9999. At 60 kW the node reads a mean of 27.761°C, a spread of 0.251 K and a probability of 0.1492, but $u^* = 1.125$ there, so the node is void: the valve is at its stop and the supply air has floated up. The saturated plant's balance, $T = T_w + Q^*(1/W + 1/G)$, gives 28.25°C on the fit (28.49°C in the stated truth), with a probability of 0.0009 by the tail and 0.0010 by the draws. The hall air is the first protected quantity, and the request for 60 kW fails it at this set point. The joint probability below is unaffected, since $T^*$ applies on the event $u^* \le 1$.

### 6.3.3 The highest set point and its back-off

**Control engineer's question.** What is the highest set point that holds it?

**General form.** Equation (6.1) inverted for its input, as equation (4.62) did for one node,

$$r_{1-\alpha} = \max\{\, r : \Pr(f(\mathbf{z}; r) \le f^* \mid \mathbf{y}) \ge 1 - \alpha \,\}, \tag{6.17}$$

under the threshold form (2.12).

**Specialized form.** Equation (6.17) with the normal form (6.3), $f = T^*$ and $\alpha = 0.05$. Since $T^*(r)$ moves one for one with $r$ and its gradient does not contain $r$, the mean at any set point is the mean at 18°C moved by $r - 18$, and the constraint inverts in closed form,

$$r_{95} = 27.5 - (T^*(18; \hat{\mathbf{z}}) - 18) - \Phi^{-1}(0.95)\sqrt{\mathbf{g}_{T^*}^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}_{T^*}}, \tag{6.18}$$

with $\Phi^{-1}(0.95) = 1.6449$; this is equation (4.66) on the state with the offset. The last term is the back-off.

**Building blocks.** The means and spreads of Section 6.3.1, at each load.

**The plant's numbers.** Table 6.5 applies equation (6.18) at four loads. Each row is a load; the columns give the mean at 18°C, the spread, the back-off and $r_{95}$ for each fit, and last the largest set point the stated truth of the record itself allows. Figure 6.2 draws the two fits at their own $r_{95}$, at 45 and 54 kW, one panel per load: the horizontal axis is the hall air at the node, the vertical axis the density, the solid curve the fit with the offset free, the dashed curve the fit with it fixed, and the vertical line the limit. Each curve leaves 0.05 of its area beyond the line, and the distance from its center to the line is its back-off.

Table 6.5: The highest set point at 0.95, equation (6.18), in the two fits, with the largest set point the stated truth allows. All columns in °C or K as marked. †: below the load's floor (Table 6.6), where the node is void and the figure describes no plant.

| | Offset free | | | | Offset fixed at 0 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $Q^*$ [kW] | Mean [°C] | sd [K] | Back-off [K] | $r_{95}$ [°C] | Mean [°C] | sd [K] | Back-off [K] | $r_{95}$ [°C] | Truth [°C] |
| 30 | 22.8521 | 0.0829 | 0.1364 | 22.5115 | 22.8523 | 0.0828 | 0.1362 | 22.5115 | 22.6276 |
| 45 | 25.3067 | 0.1559 | 0.2565 | 19.9368 | 25.2784 | 0.1242 | 0.2043 | 20.0173 | 20.0414 |
| 54 | 26.7794 | 0.2122 | 0.3491 | 18.3715 | 26.7341 | 0.1491 | 0.2452 | 18.5207 | 18.4897 |
| 60 | 27.7613 | 0.2512 | 0.4132 | 17.3255† | 27.7046 | 0.1656 | 0.2724 | 17.5230† | 17.4552† |

The curves widen and part as the load moves from the measured condition, the cancellation of equation (6.14) weakening. At 30 kW the two fits give one set point. At 45 kW the offset costs $20.0173 - 19.9368 = 0.0805$ K of set point, and both answers are inside the truth's 20.0414°C. At 54 kW the difference matters: the sensor-true fit's 18.5207°C stands above the truth's 18.4897°C, and the offset-free 18.3715°C below it and above the valve's 18.0877°C, so the true hall air would be 27.531°C at the first and 27.382°C at the second. That is one record's draw. Over 1000 fresh records of the same plant, the sensor-true $r_{95}$ passes the truth's largest set point in 49.4 % of records (se 1.6) and the offset-free one in 16.8 % (se 1.2). Neither rate is 5 %: the level is a statement given a record under the stated prior, not a frequency at one fixed truth, and this truth's offset stands 1.5 of the prior's spreads out.

At 60 kW neither set point is reachable. Both, and the truth's own 17.4552°C, lie below that load's floor, so the valve is at its stop, the node is void and the true hall air is 28.489°C at either. And no figure here settles the request, because the hall air is only one of the loop's two limits.

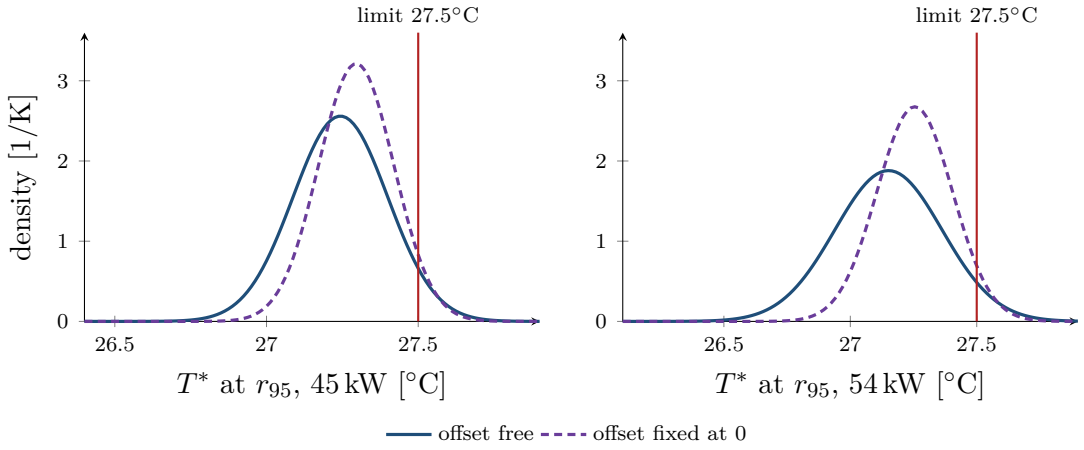



Figure 6.2: The hall air at each fit's own $r_{95}$, drawn as $\mathcal{N}(27.5 - \text{back-off},\ \text{sd}^2)$ with the back-offs and spreads of Table 6.5. Left, 45 kW: back-offs 0.2565 and 0.2043 K. Right, 54 kW, the far load in the table whose set points the valve can still reach: 0.3491 and 0.2452 K.

### 6.3.4 The valve as a protected quantity

**Control engineer's question.** Is the valve a protected quantity too?

**General form.** The chance constraint (6.1) read through the normal form (6.3), inverted for the smallest input in place of the largest, again under the threshold form (2.12).

**Specialized form.** Equations (2.12) and (6.1) with $f(\mathbf{z}; r) = u^*(r)$ of equation (6.7) and $f^* = 1$. The valve has no limit of its own on the plant's safety. It is taken as protected because a saturated loop holds nothing: at the stop the controller's row in equation (6.4) no longer applies, the supply air floats, and every answer read at the pinned $r$ is void. This is a choice of what the chance constraint is for, and it fails where a brief excursion to the stop is acceptable, which a steady-state graph cannot judge. Here the set point does not enter additively. By equation (6.13) the gradient holds both $u^*(r)$ and $r - b - T_w$, so the spread in equation (6.3) is a function of $r$,

$$\mathbf{g}_{u^*}(r)^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}_{u^*}(r) = u^*(r)^2 \left( \Sigma_{g_w g_w} - \frac{2\,\Sigma_{g_w b}}{r - \hat{b} - T_w} + \frac{\Sigma_{bb}}{(r - \hat{b} - T_w)^2} \right), \tag{6.19}$$

and the constraint is solved for $r$ by bisection with the spread recomputed at each trial, since $\Pr(u^*(r) \le 1 \mid \mathbf{y})$ rises with $r$ across the range used.

**Building blocks.** $\hat{g}_w$, $\hat{b}$ and the lower block of $\boldsymbol{\Sigma}$; equation (6.19); $\Phi$.

**The plant's numbers.** At 45 kW the spread of $u^*$ is 0.0197 at 17°C, 0.0050 at 18°C and 0.0101 at 20.5°C. It is smallest near the set point of the record, where the bracket of equation (6.19) is the combination the valve reading fixes, and it grows on both sides: the same cancellation as the hall air's, now moving with the set point rather than the load. Table 6.6 gives the result. Each row is a load; the columns give the smallest set point at 0.95, the set point at which the valve reaches its stop at the fit's mean, and the same in the stated truth. The difference between the first two columns is the valve's back-off in set point.

The floor at 60 kW has moved from Chapter 1's rated 17.75°C to 18.5524°C on the fitted coil. Lowering the set point runs into the valve, and raising it runs into the hall's limit.

Table 6.6: The smallest set point that keeps the valve inside its stop, with the mean-valued floor and the truth's floor.

| $Q^*$ [kW] | $r$ at 0.95 [°C] | Mean-valued floor [°C] | Truth's floor [°C] |
|---|---|---|---|
| 30 | 16.3968 | 16.2743 | 16.3722 |
| 45 | 17.4378 | 17.3829 | 17.4083 |
| 60 | 18.5524 | 18.4915 | 18.4444 |

### 6.3.5 The window of set points and the load that closes it

**Control engineer's questions.** Which set points hold both? At what load does no set point hold?

**General form.** The chance constraint (6.1) with the indicator of an intersection, which equation (2.12) allows for a joint event,

$$\Pr\big(f_1(\mathbf{z};r) \le f_1^*,\ f_2(\mathbf{z};r) \le f_2^* \mid \mathbf{y}\big) \ge 1-\alpha, \tag{6.20}$$

one event with one probability, not one probability per limit.

**Specialized form.** Equation (6.20) with $f_1 = T^*(r)$, $f_1^* = 27.5$°C, $f_2 = u^*(r)$ and $f_2^* = 1$. Under the normal posterior the two are jointly normal through their gradients, with the correlation

$$\rho(r) = \frac{\mathbf{g}_{T^*}^{\mathsf{T}}\,\boldsymbol{\Sigma}\,\mathbf{g}_{u^*}(r)}{\sqrt{\mathbf{g}_{T^*}^{\mathsf{T}}\,\boldsymbol{\Sigma}\,\mathbf{g}_{T^*}}\ \sqrt{\mathbf{g}_{u^*}(r)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}_{u^*}(r)}}. \tag{6.21}$$

With $z_T$ and $z_u$ the two arguments of the tail (6.2), the joint probability is the bivariate normal integral, done by deterministic quadrature on its one remaining dimension,

$$\Pr\big(T^* \le 27.5\text{°C},\ u^* \le 1 \mid \mathbf{y}\big) = \int_{-\infty}^{z_T} \varphi(a)\,\Phi\!\left(\frac{z_u - \rho\,a}{\sqrt{1-\rho^2}}\right)\mathrm{d}a. \tag{6.22}$$

The window at a load is the set of $r$ at which it reaches 0.95. The load in question is a pin of the node, so a sweep over $Q^*$ uses the same $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ and asks no more of the record. Normality near the stop is the assumption that carries this form. The valve of the holding regime has no kink of its own, so this normal meets none; the kink is in the plant past the stop, and Chapter 15 takes it up.

**Building blocks.** $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$; the gradients (6.13) and the correlation (6.21); the integral (6.22).

**The plant's numbers at two loads.** Table 6.7 evaluates the node over a grid of set points. Each row is a set point; the columns give the means and the valve's spread, the correlation, the two marginal probabilities and the joint by the state route, and the same three by the reweighted draws of Section 6.4.2, with the standard error of the joint. Figure 6.3 draws the state route at 45 kW: the horizontal axis is the set point, the vertical axis the probability, the dashed curves the two marginals, the solid curve the joint, the horizontal line the level 0.95, and the two dotted verticals the window's edges.

At 45 kW the probability of 0.95 is held for $r$ from 17.438 to 19.937°C, each edge the root of the joint at the level. The figure shows why a window exists: the valve's curve rises steeply on the left, the hall's curve falls on the right, and each binds alone at its own end. The valve's edge is steeper than the hall's because the valve's spread there, in set point, is small. The joint curve lies on one marginal or the other at every printed point, so the edges are the single-limit answers

Table 6.7: The window at two loads. The spread of $T^*$ is 0.156 K at every set point at 45 kW and 0.251 K at 60 kW. State route by the normal tails and equation (6.22); reweighted draws from a $t_5$ proposal with seed 6 and $N = 2 \times 10^6$, effective size 1 209 705. †: $u^* > 1$ at the mean, so the node is void and $\Pr(T^*)$ describes no plant; the joint is unaffected, since $T^*$ applies on the event $u^* \le 1$. A printed 1.0000 is a far tail, above 0.9999.

| | | | | | State route | | | Reweighted draws | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $r$ [°C] | $T^*$ [°C] | $u^*$ [–] | sd $u^*$ [–] | $\rho$ [–] | $\Pr(T^*)$ | $\Pr(u^*)$ | Joint | $\Pr(T^*)$ | $\Pr(u^*)$ | Joint | se |
| $Q^* = 45$ kW | | | | | | | | | | | |
| 17.00† | 24.307 | 1.130 | 0.0197 | +0.642 | 1.0000 | 0.0000 | 0.0000 | 1.0000 | 0.0000 | 0.0000 | 0.0000 |
| 17.25† | 24.557 | 1.042 | 0.0134 | +0.636 | 1.0000 | 0.0009 | 0.0009 | 1.0000 | 0.0002 | 0.0002 | 0.0000 |
| 17.50 | 24.807 | 0.966 | 0.0089 | +0.602 | 1.0000 | 0.9999 | 0.9999 | 1.0000 | 0.9997 | 0.9997 | 0.0000 |
| 17.75 | 25.057 | 0.901 | 0.0061 | +0.485 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0000 |
| 18.00 | 25.307 | 0.843 | 0.0050 | +0.225 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0000 |
| 18.50 | 25.807 | 0.749 | 0.0060 | −0.241 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0000 |
| 19.00 | 26.307 | 0.673 | 0.0076 | −0.399 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0000 |
| 19.50 | 26.807 | 0.611 | 0.0089 | −0.460 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0000 |
| 20.00 | 27.307 | 0.560 | 0.0096 | −0.490 | 0.8924 | 1.0000 | 0.8924 | 0.8991 | 1.0000 | 0.8991 | 0.0003 |
| 20.50 | 27.807 | 0.516 | 0.0101 | −0.508 | 0.0246 | 1.0000 | 0.0246 | 0.0265 | 1.0000 | 0.0265 | 0.0002 |
| $Q^* = 60$ kW | | | | | | | | | | | |
| 17.50† | 27.261 | 1.288 | 0.0119 | +0.691 | 0.8290 | 0.0000 | 0.0000 | 0.8395 | 0.0000 | 0.0000 | 0.0000 |
| 18.00† | 27.761 | 1.125 | 0.0067 | +0.189 | 0.1492 | 0.0000 | 0.0000 | 0.1592 | 0.0000 | 0.0000 | 0.0000 |
| 18.25† | 28.011 | 1.058 | 0.0070 | −0.157 | 0.0209 | 0.0000 | 0.0000 | 0.0229 | 0.0000 | 0.0000 | 0.0000 |
| 18.50 | 28.261 | 0.998 | 0.0080 | −0.368 | 0.0012 | 0.5945 | 0.0002 | 0.0014 | 0.5902 | 0.0002 | 0.0000 |
| 18.75 | 28.511 | 0.945 | 0.0091 | −0.481 | 0.0000 | 1.0000 | 0.0000 | 0.0000 | 1.0000 | 0.0000 | 0.0000 |
| 19.00 | 28.761 | 0.897 | 0.0102 | −0.545 | 0.0000 | 1.0000 | 0.0000 | 0.0000 | 1.0000 | 0.0000 | 0.0000 |
| 19.50 | 29.261 | 0.815 | 0.0118 | −0.611 | 0.0000 | 1.0000 | 0.0000 | 0.0000 | 1.0000 | 0.0000 | 0.0000 |

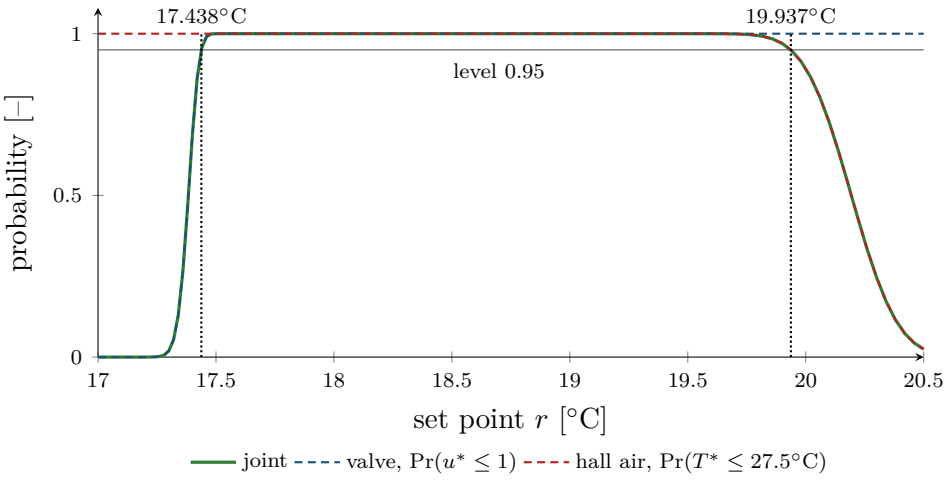



Figure 6.3: The joint chance constraint at $Q^* = 45$ kW, closed loop, by the state route: the tails (6.2) and the integral (6.22) with the fit of Section 6.2.2. The window is where the solid curve stands above the level.

17.4378 and 19.9368°C of Tables 6.6 and 6.5, to the three decimals printed. At 60 kW there is no such range: the largest joint probability is 0.0004, at $r = 18.556$°C, and the marginals in Table 6.7 show why. Where the valve holds, the hall does not.

The correlation changes sign across the window, from +0.642 at 17°C to −0.508 at 20.5°C, and yet it hardly matters. The reason is where the two limits bind. They bind at opposite ends of the window, and at an end one marginal is close to one, so the joint equals the other whatever the correlation. Only near closure are both marginals far from one at the same set point, so that is the test: at the best $r$ the joint against the product of the marginals is 0.9493 against 0.9494 at 55 kW and 0.7182 against 0.7203 at 56 kW. Far past closure the two part as a ratio, 0.000365 against 0.000545 at 60 kW, a factor of 1.49, on probabilities no decision reads.

**The load that closes it.** Table 6.8 repeats the root-finding at nine loads. Each row is a load; the columns give the largest joint probability, the set point at which it is reached, the window's two edges and its width. Figure 6.4 draws the same result: the horizontal axis is the load and the vertical axis the set point, the upper solid curve the window's top edge, the lower solid curve its bottom edge, the dashed chords the two single-limit answers continued from 54 to 60 kW, and the dotted lines the same two edges in the stated truth, which meet at the dotted vertical.

Table 6.8: The window of set points against the load, closed loop. Below the level the window is empty and the best probability is the answer.

| $Q^*$ [kW] | Best joint [–] | at $r$ [°C] | Lower edge [°C] | Upper edge [°C] | Width [K] |
|---|---|---|---|---|---|
| 45 | >0.9999 | – | 17.438 | 19.937 | 2.499 |
| 48 | >0.9999 | – | 17.649 | 19.416 | 1.767 |
| 50 | >0.9999 | – | 17.792 | 19.068 | 1.276 |
| 52 | >0.9999 | – | 17.938 | 18.720 | 0.782 |
| 54 | 0.9969 | 18.133 | 18.088 | 18.372 | 0.283 |
| 55 | 0.9493 | 18.190 | empty | | – |
| 56 | 0.7182 | 18.252 | empty | | – |
| 58 | 0.0730 | 18.393 | empty | | – |
| 60 | 0.0004 | 18.556 | empty | | – |

The window narrows from 2.499 K at 45 kW to 0.283 K at 54 kW, and it is empty at 55 kW, where the best probability is 0.9493, at 56 kW with 0.7182, at 58 kW with 0.0730 and at 60 kW with 0.0004. So the 60 kW request of Chapter 1 fails for a reason that chapter could only point to: no set point holds it. At that load the coil's floor has risen to 18.5524°C, above the 17.3255°C that the hall's form returns there, a void reading. Root-finding the best joint probability at the level places the closure at 54.99 kW by the state route; the two single-limit curves cross at 55.13 kW. The stated truth's own limit is 55.90 kW, a figure with no band, so the two are different quantities. The window is a reading at one pinned load with a normal posterior near the stop, and the load is pinned, not forecast.

## 6.4 The checks (step 5)

Four tests stand against the answers: the limit of exact physical rows, the exact posterior by reweighted draws, the prior that the data do not correct, and the stated truth of the record.

### 6.4.1 The limit of exact rows

The reduced state assumes the rows (6.4) hold exactly. Table 6.9 refits the seven-entry full state with those rows at a finite tolerance, the valve's at $0.01\,\sigma_p$. Each column is a tolerance,

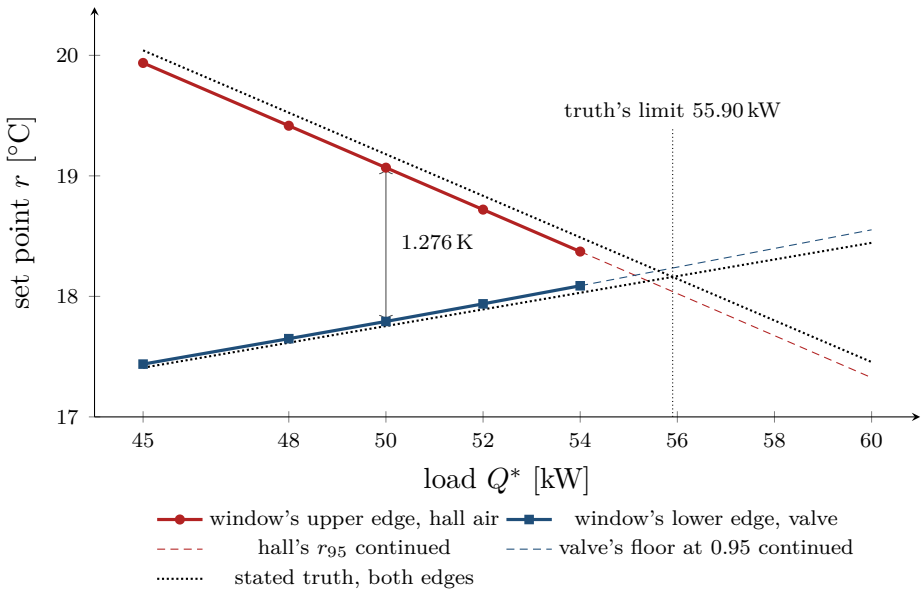



Figure 6.4: The window of set points against the load, from Table 6.8, with the points at 60 kW from Tables 6.5 and 6.6; the dashed chords join those points and are drawn for the eye, and past the floor the hall's $r_{95}$ is a void reading. The truth's edges are straight lines: floor $r = T_w + b + Q^*/W$, limit $r = 27.5 + b - Q^*/G$.

the last the reduced state, and the rows give the two entries the chapter's result rests on. The full state reaches the reduced state's values by $\sigma_p = 0.001$, within the spread's fourth decimal. The Gauss–Newton spreads without the curvature term of equation (6.11) are 0.158, 0.04171, 0.04797 and 0.18793 for $Q$, $g$, $g_w$ and $b$, against the 0.15798, 0.04210, 0.04846 and 0.18986 reported; this chapter's curvature is small but no longer negligible in the third figure.

Table 6.9: The full state against the reduced state as the physical rows' tolerance falls. Standard deviations in parentheses.

| Entry [unit] | $\sigma_p = 0.100$ | $\sigma_p = 0.010$ | $\sigma_p = 0.001$ | Reduced state |
|---|---|---|---|---|
| $g_w$ [–] | −0.16442 (0.05361) | −0.16761 (0.04851) | −0.16764 (0.04844) | −0.16764 (0.04846) |
| $b$ [K] | 0.05575 (0.19005) | 0.05707 (0.18987) | 0.05708 (0.18982) | 0.05708 (0.18986) |

### 6.4.2 The exact posterior by reweighted draws

The second route draws $N = 2 \times 10^6$ states from a multivariate $t$ with five degrees of freedom about $\hat{\mathbf{z}}$, on the Laplace scale, seed 6, weights each by the exact posterior over that proposal, and pushes every draw through equation (6.7) without linearizing. Table 6.7 carries its probabilities beside the state route's, like with like, and its standard errors come from the weights. The proposal is a choice this posterior forces. Chapter 5's normal proposal, at this $N$, gives an effective size of 308 361, 109 876 and 10 021 on seeds 6, 7 and 8, with standard errors that themselves cannot be trusted. Since $\mathrm{corr}(g_w, b) = 0.992$, the normal is a thin ridge, the rows are curved along it through $1/(r - b - T_w)$, and the exact posterior's tail along the ridge outruns the normal's, so a rare draw takes a very large weight. The $t_5$'s heavier tail covers it: its effective size is 1 209 705, 1 209 453 and 1 210 468 on the same seeds.

The largest gap in the joint probability on the grid is at 20°C and 45 kW, 0.8924 by the state route against 0.8991 by the draws, with a standard error of 0.0003, so the gap is real and not sampling noise. Its cause is not the curvature of $T^*$ in $g$: that alone, on the Laplace

normal, moves the probability the other way, to 0.8821. The exact posterior with $T^*$ kept linear gives 0.9091, and the two together give 0.8991, so the posterior's non-normality outweighs the curvature. Off the grid the gap peaks at +0.016 near $r = 20.2$°C (se 0.0005). The values at 20°C lie on the same side of 0.95, and nothing in the window's verdict changes: by the draws the window at 45 kW is 17.440 to 19.942°C, moving its edges by +0.002 and +0.005 K. The gap is larger on the void rows, 0.8290 against 0.8395 at 17.5°C and 60 kW, which describe no plant. At the valve's edge at 45 kW the two routes read 0.9999 and 0.9997 at 17.5°C, and 0.0009 and 0.0002 at 17.25°C.

### 6.4.3 Sensitivity to the prior on the offset

The prior on $b$ is a choice, and since the readings barely see the offset, nothing in the record corrects it. Table 6.10 refits with five prior spreads. Each row is a prior; the columns give the offset and the coil with their spreads, $r_{95}$ at 45 kW and the window at 45 kW.

Table 6.10: Sensitivity to the prior on the offset, at 45 kW. The third row is the stated choice. The truth's largest set point is 20.0414°C and its floor 17.4083°C.

| Prior sd $b$ [K] | $\hat{b}$ [K] | $\sqrt{\Sigma_{bb}}$ [K] | $\hat{g}_w$ [–] | $\sqrt{\Sigma_{g_w g_w}}$ [–] | $r_{95}$ [°C] | Window [°C] |
|---|---|---|---|---|---|---|
| 0.05 | 0.0039 | 0.0498 | −0.1810 | 0.0138 | 20.0113 | 17.410 to 20.011 |
| 0.10 | 0.0154 | 0.0987 | −0.1781 | 0.0254 | 19.9943 | 17.417 to 19.994 |
| **0.20** | 0.0571 | 0.1899 | −0.1676 | 0.0485 | 19.9368 | 17.438 to 19.937 |
| 0.40 | 0.1748 | 0.3296 | −0.1374 | 0.0863 | 19.7968 | 17.481 to 19.797 |
| 0.80 | 0.3491 | 0.4521 | −0.0908 | 0.1238 | 19.6257 | 17.528 to 19.626 |

The estimate of the offset runs from 0.004 to 0.349 K, tracking its prior, and its spread stays near the prior's until the prior is wide. That the prior sets the offset's estimate is a further sign the data do not see it, and no reading of this record will say which row of the table is right. The set point the answer allows falls with the prior's width, from 20.011 to 19.626°C, and the window goes from 17.410–20.011 to 17.528–19.626°C. On this record every row stays inside the truth's own edges at 45 kW, but the width of the prior is a quantity the commissioning record must state, because it moves the answer and the data will not move it back.

### 6.4.4 The answers against the stated truth

The record is constructed from a stated truth, and that truth holds an offset of +0.30 K on the discharge sensor. The loop's error of +0.0032 K shows none of it. Table 6.11 compares the fit with the truth: each row is an entry or a node, the columns the fit, its spread, the truth and the distance in the fit's standard deviations. The offset's own distance is +1.28 standard deviations, the coil's +1.40 and the path's −1.24, and they share one sign pattern because they are one error along one ridge.

Figure 6.5 draws the posterior in the plane of $(g_w, b)$. The horizontal axis is the offset, the vertical axis the coil's log-derate; the two ellipses are the one- and two-standard-deviation contours of the Laplace normal, the dot is the MAP state, the square with its short bar is Chapter 5's fit with its spread, and the cross is the stated truth, with dotted lines to the axes.

The figure explains the 13.77. Chapter 5's fit is the slice of this posterior at $b = 0$: it lies on the ridge, with the ridge's narrow width for its spread. The truth lies farther along the same ridge, inside the two-standard-deviation contour. That fit was not misread; it was certain about one point of a line the data leave open, and both points meet the readings. It also says why the operating answers barely moved. The node at the record's load and set point reads across the ridge, so which point along it holds hardly matters there.

Table 6.11: The fit with the offset free against the stated truth, the entries with their distance in the fit's standard deviations and the node at 45 kW beside them.

| Entry or node [unit] | Fit | sd | Truth | Truth minus fit [sd] |
|---|---|---|---|---|
| $Q$ [kW] | 30.05809 | 0.15798 | 30.25634 | +1.25 |
| $g$ [–] | 0.01833 | 0.04210 | −0.03390 | −1.24 |
| $g_w$ [–] | −0.16764 | 0.04846 | −0.10000 | +1.40 |
| $b$ [K] | 0.05708 | 0.18986 | 0.30000 | +1.28 |
| $T_c$ [°C] | 17.94292 | 0.18986 | 17.69368 | −1.31 |
| $T^*$ at 18°C, 45 kW [°C] | 25.3067 | 0.1559 | 25.4586 | – |
| $u^*$ at 18°C, 45 kW [–] | 0.8435 | 0.0050 | 0.8401 | – |

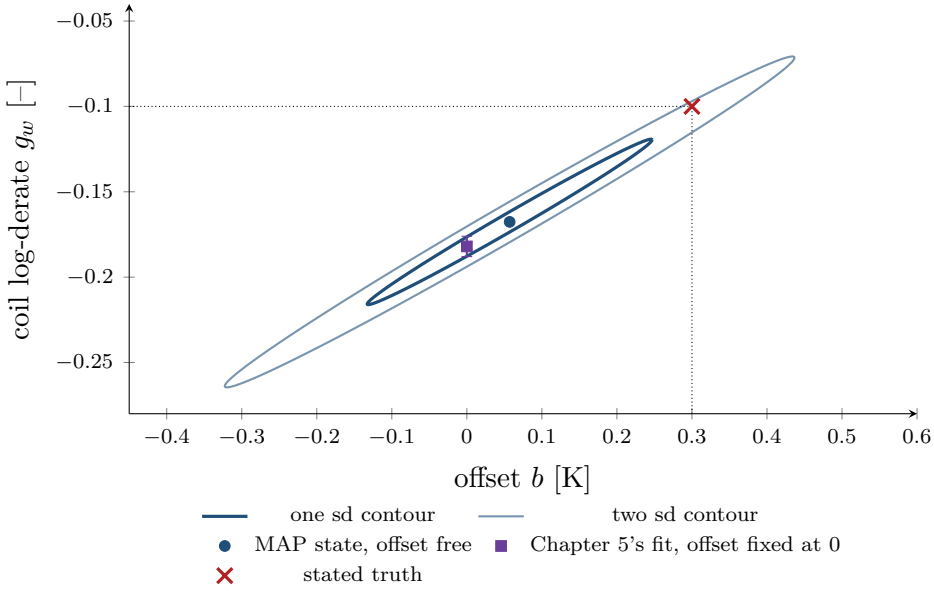



Figure 6.5: The posterior of $(g_w, b)$, the Laplace normal with spreads 0.04846 and 0.18986 and correlation 0.99242. The contours are the points $(\hat{b} + k\sqrt{\Sigma_{bb}}\cos t,\ \hat{g}_w + k\sqrt{\Sigma_{g_w g_w}}(\rho\cos t + \sqrt{1-\rho^2}\sin t))$ for $k = 1, 2$.

## 6.5 The report (step 6)

The window of set points closes at 54.99 kW. The request for 45 kW can be met at any set point from 17.438 to 19.937°C, 18°C among them, and the request for 60 kW cannot be met at any set point, with a best joint probability of 0.0004. In the constructed record the discharge sensor reads +0.30 K high, invisible in the loop's error of +0.0032 K. The Freeing the offset moved the coil's estimate by more than Chapter 5's spread could allow, and it moved the 45 kW set point by 0.08 K.

### 6.5.1 The answers and their bands

Table 6.12 gives the answers in the order of Table 6.1. Each row is a question; the second column gives the state-route answer, with the reweighted draws beside it where they were computed; the third gives the band, here a standard deviation or the range over the prior widths of Table 6.10; the last gives the stated truth.

Table 6.12: The report for the closed-loop hall, offset free, $\alpha = 0.05$. Answers by the state route; draws from a $t_5$ proposal with seed 6, $N = 2 \times 10^6$, effective size 1 209 705. The draws were computed for the probabilities on the grid of Table 6.7 and the edges at 45 kW, not for the moments.

| Control engineer's question | Answer | Band | Stated truth |
|---|---|---|---|
| What does an offset the loop cannot see do to the plant? | $\hat{b} = 0.0571$ K; $T^*$ at 18°C, 45 kW: 25.307°C; $u^*$: 0.8435 | sd 0.1899 K, information 2.74 $K^{-2}$ against the prior's 25, all through the coil's and path's priors; sd 0.156 K; sd 0.0050 | $b = 0.30$ K; 25.459°C; 0.8401 |
| What must stay inside its limit, with what probability? | $\Pr(T^* \le 27.5°C)$ at 18°C: >0.9999 at 45 kW; at 60 kW the node is void (0.1492), the saturated plant 0.0009 (draws 1.0000; 0.0010) | – | saturated plant 28.49°C at 60 kW |
| What is the highest set point that holds it? | $r_{95} = 19.937$°C at 45 kW; 18.372°C at 54 kW; at 60 kW void | back-off 0.2565; 0.3491 K. Over the prior widths at 45 kW: 19.626 to 20.011°C | 20.041°C; 18.490°C |
| Is the valve a protected quantity too? | smallest $r$ at 0.95: 17.438°C at 45 kW; 18.552°C at 60 kW | spread of $u^*$ at 45 kW: 0.0197 at 17°C, 0.0050 at 18°C, 0.0101 at 20.5°C | 17.408°C; 18.444°C |
| Which set points hold both? | at 45 kW, 17.438 to 19.937°C (draws 17.440 to 19.942°C). At 20°C the joint is 0.8924 (draws 0.8991, se 0.0003) | over the prior widths, 17.410–20.011 to 17.528–19.626°C | 17.408 to 20.041°C |
| At what load does no set point hold? | the window closes at 54.99 kW; width 0.283 K at 54 kW; best joint 0.9493 at 55 kW, 0.0004 at 60 kW | – | 55.90 kW |

### 6.5.2 The entry in the commissioning log

The entry states what is protected and at what level: the hall air against 27.5°C and the valve against its stop, jointly at 0.95, the valve because a saturated loop holds nothing. At that level the record supports 45 kW at any set point from 17.438 to 19.937°C and does not support 60 kW at any set point. It says the sensor is not cleared: the discharge sensor's offset is estimated at 0.06 K with a spread of 0.19 K, which is almost its calibration tolerance, the loop's error says nothing about it, and the estimate follows the prior. It says that the fit's temperature and valve

at the present set point do not depend much on that, and that a coil figure does. Where in the window the set point should stand is left open: the entry does not choose, and what a tighter posterior is worth is the question of Chapter 16. The verdict rests on one record's draw of its sensors' errors.

### 6.5.3 Common mistakes

Table 6.13 lists five mistakes this chapter's plant makes easy. Each row names the mistake, its form on this hall and the correction.

Table 6.13: Five mistakes in protecting a quantity under feedback.

| Mistake | On this hall | Correction |
|---|---|---|
| Reading a small loop error as a sound sensor | $+0.0032\,\mathrm{K}$ on the trend, with a $+0.30\,\mathrm{K}$ offset in the record | the discharge reading has a zero line on the state; free the offset and give it a prior |
| Taking a passed misfit test as clearing a choice | $P = 0.388$ and $P = 0.406$, one fit 13.77 sd from the coil | the misfit tests the readings against the rows; compare the choices on what they do to the answers |
| Adding the offset's whole spread to the band | $0.0961 + 0.0360$ without the cross term $-0.1078\,\mathrm{K}^2$ | carry $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ with its cross terms; the spread depends on the distance from the measured condition |
| Using one back-off at every set point | the valve's spread is 0.0050 at 18°C and 0.0197 at 17°C | recompute the gradient at each $r$ wherever $r$ is not additive |
| Protecting the temperature alone | $r_{95} = 17.33$°C at 60 kW, below the valve's 18.55°C, where the node is void | name every limit the loop has; the set point must hold them in one event |

## 6.6 The next measurement (step 7)

The band's largest open question is an entry the readings cannot reach. A further reading is one more row of the residual vector, and what it would remove is known before its value is.

### 6.6.1 The preposterior of the candidates

A candidate with Jacobian line $\mathbf{J}_a$ gives the preposterior of equation (4.73),

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1}, \tag{6.23}$$

which does not contain the value read. In the columns of the state, a reference discharge sensor reads the air directly, $\mathbf{J}_a = (0,\ 0,\ 0,\ -1)/0.05$, the one line in this chapter with no column but the offset's. A permanent second discharge sensor has the same line, divided by its own calibration tolerance of 0.2 K combined with its noise over the forty steps, since its offset is a second offset. A second hall sensor repeats the hall reading's line. A water-side coil duty meter of 1 kW and a load meter of 0.2 kW, each averaged over the forty steps, both read the load, on the line $(1,\ 0,\ 0,\ 0)$ over their standard deviations.

### 6.6.2 The candidates

Table 6.14 ranks them at $r = 18$°C and 45 kW, from the present spreads 0.1899 K, 0.1559 K and 0.0050. Each row is a candidate; the columns give the spreads of the offset, of the hall air and of the valve after it by equation (6.23), and $r_{95}$ with the mean held where it is. Figure 6.6 draws what each removes from the first two spreads: the horizontal axis is the reduction in K, the bars the two spreads for each candidate.

Table 6.14: The candidates at $r = 18°\text{C}$, 45 kW, with $r_{95}$ at the present mean. At present $r_{95} = 19.9368°\text{C}$.

| Candidate | sd $b$ [K] | sd $T^*$ [K] | sd $u^*$ [–] | $r_{95}$ [°C] |
|---|---|---|---|---|
| Reference discharge sensor, 0.05 K | 0.0484 | 0.1266 | 0.0050 | 19.9850 |
| Second hall sensor, the same 20 minutes | 0.1899 | 0.1319 | 0.0050 | 19.9763 |
| Permanent second discharge sensor, its own offset 0.2 K | 0.1379 | 0.1419 | 0.0050 | 19.9599 |
| Load meter, 0.2 kW, averaged | 0.1899 | 0.1511 | 0.0025 | 19.9448 |
| Water-side coil duty meter, 1 kW, averaged | 0.1899 | 0.1534 | 0.0039 | 19.9409 |

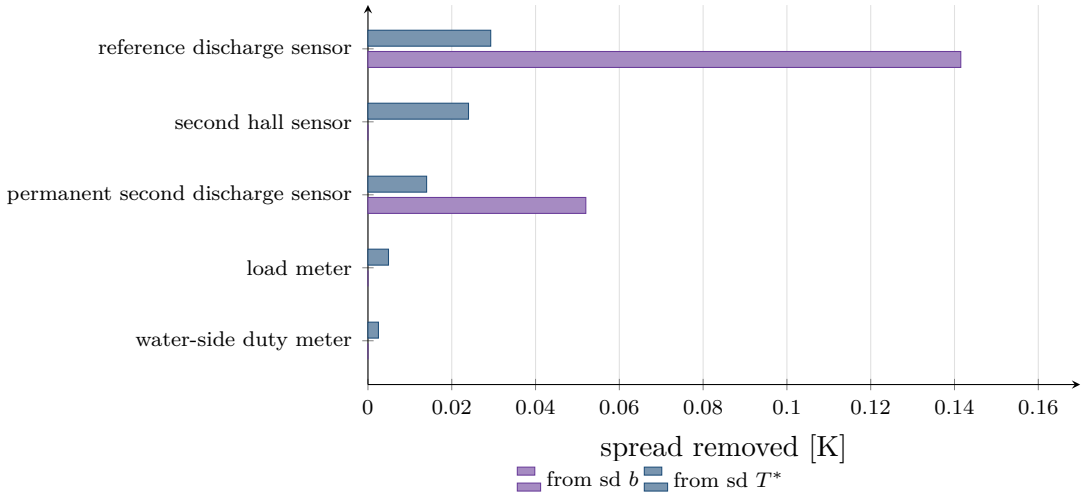



Figure 6.6: The spread each candidate of Table 6.14 removes: the present spread less the spread after it, by the preposterior (6.23).

Chapter 5's ranking reverses for the discharge sensor. There, with the sensor taken as true, a reference removed nothing from the spread of $T^*$; here, it does the most, and only a sensor at the discharge sees the offset at all, as the column of sd $b$ shows. A permanent sensor of the same tolerance only averages two offsets, so a brief reference does more than a permanent second sensor. The second hall sensor removes nearly as much from $T^*$ without touching the offset, since it narrows the combination the node reads. The two load instruments act on the valve's spread through the load, and do little for the hall air.

**A reference reading.** The candidates hold the mean still, and a real reading moves it. As a constructed example, take the reference at its 0.05 K reading 0.30 K, the record's truth. This reading is constructed for the illustration and no sensor returned it. The refit gives $\hat{b} = 0.2841$ (0.0483), $\hat{g} = -0.02676$ (0.01874) and $\hat{g}_w = -0.10841$ (0.01429), with the truth at $-0.38$ standard deviations on the path and $+0.59$ on the coil. The window at 45 kW becomes 17.452 to 19.871°C, against 17.438 to 19.937°C. So even then the window moves little, and at the top it falls, a narrower posterior with a lower set point. The reference's value lies in the parameters, for the diagnosis of later chapters, more than in today's set point.

### 6.6.3 The recommendation

The next measurement is a reference discharge sensor, brought once. It takes the offset's spread from 0.1899 to 0.0484 K and it is the only candidate that settles what the coil and the path are. The gain in the set point is small: $r_{95}$ moves from 19.9368 to 19.9850°C with the mean held. As the permanent instrument, a second discharge sensor of the same tolerance is the weaker request, and a second hall sensor gives the band nearly as much, in every record. What a tighter posterior is worth in set point, over the values it might return, is the question of Chapter 16 and is not settled here. Chapter 26 uses the coil and the path this sensor would recover.

### 6.6.4 What the analysis showed

The trend said the loop was holding with no error, and a fit that took the sensor as true passed its test; the analysis finds a record whose offset the loop cannot see and whose coil that fit had badly wrong, yet whose answers on the plant's limits, at this prior, hardly moved. The readings fix the combinations the node reads, so the offset is cheap at the condition the hall was measured at and grows dearer with distance from it: at 54 kW the sensor-true $r_{95}$ breaks the limit in 49.4 % of fresh records and the offset-free one in 16.8 %. The commissioning record must therefore carry the offset even though it cannot estimate it, and state the prior it chose. The hall's real limit is not the temperature alone. A loop has a hall limit above and a valve stop below, and the two together answer the request: the set point has room at the lower load, none at the higher, and the hall's capacity under this loop ends where the two meet.

### 6.6.5 Limits of the model

**The normal posterior near the stop.** Equation (6.22) treats $u^*$ as normal on both sides of 1, where the plant has a kink. The window's lower edge is a reading of that normal, and Chapter 15 takes up the kink.

**Steady state.** Every node is a steady operating point. The window says nothing of how the loop gets there, or of whether it is stable on the way, which is Chapter 8's question.

**The valve as protected.** The valve is a limit because a saturated loop holds nothing. A plant that accepts short visits to the stop would ask a different constraint.

**One pinned load.** A window applies at one load, and nothing here forecasts the load. A load that moves within the hour needs Chapter 19's treatment.

**The prior on the offset.** The 0.2 K is a stated choice that the data do not correct, and Table 6.10 shows how far it moves the answers. One constant offset is assumed for every set point; a sensor whose error depends on its temperature would not read the same at the other end of the window.

**The constructed record.** The readings are constructed from a stated truth, and the truth's agreement with the fit tests the method on its own model, not the model against a real hall. Chapter 8 asks whether this tuning is stable given the same band.

## 6.7 What to retain

**Definition 6.1** (Chance constraint)**.** A chance constraint requires a limit on a function of the state, at an input pinned at a scenario node, to hold with a stated probability, $\Pr(f(\mathbf{z}; r) \le f^* \mid \mathbf{y}) \ge 1 - \alpha$. For a normal posterior it reads $f(\hat{\mathbf{z}}; r) + \Phi^{-1}(1-\alpha)\sqrt{\mathbf{g}(r)^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}(r)} \le f^*$, and the second term is the back-off, valid while $f$ is close to normal, with the gradient recomputed at each $r$ unless $r$ enters additively. Several limits held together are one event with one probability.

Three conclusions carry beyond the hall. A controller holds its reading, so a sensor offset is an entry of the state, invisible in the loop's error, and it passes into the protected quantity in full, with the data cancelling most of it only near the condition they were taken at. A loop has at least two protected quantities, and they bind at opposite ends of a window of set points, so correlation hardly matters, even at closure. And an entry the readings cannot see is carried by its prior, whose width is then part of the answer and has to be stated.

Table 6.15 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 6.15: The terms of the chapter and where each is defined. Units for the closed-loop hall.

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| Chance constraint | eq. (6.1) | – | a limit held with a stated probability | 6.1.1 |
| Back-off | $\Phi^{-1}(1-\alpha)\sqrt{\mathbf{g}(r)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(r)}$, eq. (6.3) | K, – | distance the mean stands inside the limit | 6.1.1 |
| Controller's row | $T_c + b = r$ | K | holds the reading, not the air | 6.1.2 |
| Reduced state | $(Q, g, g_w, b)$ | kW, –, K | free entries with the offset freed | 6.1.2 |
| Scenario node | eq. (6.7) | °C, – | closed loop, $r$ and $Q^*$ pinned | 6.1.2 |
| Zero line of the loop's reading | $y_c - r$ | K | the loop's error holds no entry | 6.1.4 |
| Split of the gradient | eq. (6.14), $\lambda = Q^*/\hat{Q}$ | K | what the readings cancel, what is left | 6.3.1 |
| Highest set point | eq. (6.18) | °C | the hall air's limit inverted | 6.3.3 |
| Valve's spread | eq. (6.19) | – | depends on $r$ | 6.3.4 |
| Joint chance constraint | eq. (6.20), eq. (6.22) | – | both limits in one event | 6.3.5 |
| Window of set points | $\{r$ : eq. (6.20) holds$\}$ | °C | range holding both at one load | 6.3.5 |
| Effective sample size | $1/\sum_i w_i^2$, normalized weights | – | what reweighted draws are worth | 6.4.2 |
| Preposterior | eq. (6.23) | mixed | the band after a candidate row | 6.6.1 |

## 6.8 Exercises

**Exercise 6.1.** *Formulation.* A pump's speed controller holds the discharge pressure reading of a chilled-water loop at a set point $r$. Name the protected quantities and write the joint chance constraint, with the entry a sensor offset adds. (The far branch's differential pressure against its required minimum, and the pump's speed against its top stop, because a saturated loop holds nothing. With offset $b$ the true discharge is $r - b$, and with $p^*(r)$ the branch pressure and $s^*(r)$ the speed at a pinned demand, $\Pr(p^*(r) \geq p_{\min},\ s^*(r) \leq 1 \mid \mathbf{y}) \geq 1 - \alpha$; written in the form of equation (6.20), $f_1 = -p^*$ and $f_1^* = -p_{\min}$.)

**Exercise 6.2.** *Derivation.* Derive the back-off from the general form and give its value for the hall air at 45 kW. (By the delta method $f \mid \mathbf{y} \approx \mathcal{N}(f(\hat{\mathbf{z}}; r),\ \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g})$, so equation (6.1) reads $\Phi((f^* - f(\hat{\mathbf{z}}; r))/\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}) \geq 1-\alpha$. Applying $\Phi^{-1}$ gives $f(\hat{\mathbf{z}}; r) + \Phi^{-1}(1-\alpha)\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}} \leq f^*$. With $\Phi^{-1}(0.95) = 1.6449$ and spread 0.1559 K the back-off is 0.2565 K.)

**Exercise 6.3.** *Derivation.* Show why $r$ leaves the hall's gradient and stays in the valve's. $(T^*(r) = r - b + Q^*/G$ is a sum with $r$ in one term, so $\partial^2 T^*/\partial r\,\partial \mathbf{z} = \mathbf{0}$ and $\mathbf{g}_{T^*} = (0,\ -Q^*/G,\ 0,\ -1)^{\mathsf{T}}$. $u^*(r) = Q^*/(W(r - b - T_w))$ has $r$ in a denominator with $b$, so $\mathbf{g}_{u^*} = u^*(r)\,(0,\ 0,\ -1,\ 1/(r - b - T_w))^{\mathsf{T}}$ contains $r$ twice. The valve's spread at 45 kW is 0.0197 at 17°C, 0.0050 at 18°C and 0.0101 at 20.5°C.)

**Exercise 6.4.** *Application.* Find the window of set points at 50 kW. (The upper edge is $r_{95}$ of equation (6.18) with $Q^* = 50$ kW in the mean and in $\mathbf{g}_{T^*}$. The lower edge solves $\Pr(u^*(r) \leq 1 \mid \mathbf{y}) = 0.95$ with the spread of equation (6.19) recomputed at each trial $r$. Between them the joint (6.22) is above the level, and it is >0.9999 at its best. Root-finding each edge gives 17.792 to 19.068°C, a width of 1.276 K.)

**Exercise 6.5.** *Application.* At 60 kW compare the joint probability with the product of the marginals at the best set point, and say what the comparison means for the two limits. (At $r = 18.556$°C the product is 0.000545 and the joint is 0.000365, a ratio of 1.49, so at this load the two do not agree. That load is far past closure, where neither holds where the other does, and the answer in both forms is that no set point holds it. The test of the correlation is near closure, where both marginals are far from one: 0.9493 against 0.9494 at 55 kW and 0.7182 against 0.7203 at 56 kW. The two limits bind at different ends of the window, so the correlation hardly enters.)

# Chapter 7

# The Control Problem on the Graph

## 7.1 One problem under every controller

The commissioning engineer has four sheets on one loop. The vendor's sheet gives the PI row of Chapter 5, $K_p = 0.10$ per kelvin and $K_i = 0.002$ per kelvin second, with no reason attached. Chapter 5's fit gives a process gain of 7.0996 K per unit of valve. Chapter 6's sheet gives a limit of 27.5°C and, at 45 kW, a spread of 0.1559 K with a back-off of 0.2565 K. On top of these lies an offer of a model predictive controller, with a horizon, weights and a mathematics of its own. Each sheet was written by someone who did not need the others.

The engineer would conclude that these are separate subjects: estimate first, tune the loop second, and put a margin on the set point last. On that reading the spread of 0.1559 K belongs in the report and never reaches the valve, because the PI row takes a reading and has no input for a band. Nothing on the four sheets tests that reading. The question of this chapter is

*whether the sheets answer one question or several, and when the band changes what the valve should do.*

**Jargon: policy, belief, coupling, restriction, Bellman recursion, certainty equivalence, separation, fixed sequence, recourse, optimism, caution, dual effect, probing**

A *policy* is a rule that gives the next move from what is known at the time of the move, whatever that turns out to be. The *belief* is what is known about the state after the readings so far; this book carries it as the fit's pair $(\hat{\mathbf{z}}_k, \mathbf{\Sigma}_k)$. A *coupling* is a way in which the parts of the problem act on each other, and a *restriction* is the problem with its policy class, horizon, plant class or objective narrowed, which gives up some of them. The *Bellman recursion* solves a problem over policies backward from the last move. *Certainty equivalence* is the use of the estimate as if it were the truth when choosing the move, and *separation* is the case in which doing so is optimal, so that the fit and the decision can be done one after the other. A *fixed sequence* is a list of moves settled now and carried out whatever is read. *Recourse* is a later move that responds to a later reading. *Optimism* is the error of minimizing over future noise instead of averaging over it. *Caution* is a move shrunk because the gain it acts through is uncertain. The *dual effect* is the fact that a move changes what will be known afterward, and *probing* is a move made for that reason.

**Key idea**

A controller is a partial answer to one problem: choose a policy that minimizes the expected consequence, with the physics, the readings, the fit and the limits as its constraints. The rows on the past are conditioned on what was read, and the same rows on the future are averaged over what has not happened. Most later chapters are named restrictions of this problem and say which couplings they keep; the rest are marked.

### 7.1.1 The problem

Chapter 2 wrote a decision as the expected consequence of an action, $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$ with $a^* = \arg\min_a \mathcal{R}(a)$, equation (2.23). Chapter 6 chose one action of this kind, a set point, under the chance constraint (6.1). A loop needs three changes to that form. There are many moves in time, not one action. The later moves are made after readings that do not yet exist, so what is chosen now is not a list of moves but a rule. And each move alters the plant that produces those readings. Written with all three, the action $a$ becomes a policy $\pi = (\pi_0, \pi_1, \ldots)$, and equation (2.23) becomes

$$
\begin{array}{lll}
\min\limits_{\pi} & \mathbb{E}\Big[\sum\limits_{k} \ell(\mathbf{z}_k, \mathbf{u}_k)\Big] & \\
\text{subject to} & F(\mathbf{z}_{k+1}, \mathbf{z}_k, \mathbf{u}_k) = \mathbf{w}_k & \text{physics, the move in its rows} \\
& \mathbf{y}_k = h(\mathbf{z}_k, \mathbf{u}_{k-1}) + \boldsymbol{\varepsilon}_k & \text{readings of the moved plant} \\
& (\hat{\mathbf{z}}_k, \boldsymbol{\Sigma}_k) = \mathrm{Fit}(\hat{\mathbf{z}}_{k-1}, \boldsymbol{\Sigma}_{k-1}, \mathbf{u}_{k-1}, \mathbf{y}_k) & \text{what the policy may see} \\
& \mathbf{u}_k = \pi_k(\hat{\mathbf{z}}_k, \boldsymbol{\Sigma}_k) & \text{the policy closes the loop} \\
& \Pr\big(g(\mathbf{z}_k, \mathbf{u}_k) \le 0 \text{ for all } k \mid \mathbf{y}_{\text{now}}\big) \ge 1 - \alpha & \text{the limits, on the belief}
\end{array}
\tag{7.1}
$$

Here $\mathbf{z}_k$ is the state at controller step $k$, $\mathbf{u}_k$ the move made after the reading $\mathbf{y}_k$, so that $\mathbf{u}_k$ depends on $\mathbf{y}_{0:k}$ and the reading depends on the move before it, $\mathbf{u}_{k-1}$; $\ell$ is the consequence of one step, $\mathbf{w}_k$ the plant's noise over the step and $\boldsymbol{\varepsilon}_k$ the sensors' noise. $\mathbf{y}_{\text{now}}$ is what has been read up to the moment of deciding, and the expectation is over the belief at that moment and all later $\mathbf{w}$ and $\boldsymbol{\varepsilon}$. The set point stays $r$; a row always carries its argument, $r(\mathbf{z})$.

Two lines of equation (7.1) need a word. The last is the joint form of equation (6.1), with the limits written $g \le 0$ and one event over the horizon. This joint form, taken once at the moment of deciding, is the definition, and the words around it matter. A limit held separately at each step, at the same $\alpha$, is not a restriction of it but a relaxation: the joint event at 0.95 implies each step's event at 0.95, so the per-step form asks for less. A limit held after every reading asks for more, and that one is a restriction; the example of Section 7.3.3 limits one step and measures only this reading-by-reading gap. The problem is stated at one now; a controller run in time restates it at every step, which replaces the joint guarantee by the conditional one of Section 7.3.3, a restriction.

The line with Fit is this book's fit, the pair of a mean and a covariance. The fit restricts the policy's information: $\mathbf{u}_k$ may depend on the record only through $(\hat{\mathbf{z}}_k, \boldsymbol{\Sigma}_k)$, while the expectation stays under the true posterior $p(\mathbf{z}_k \mid \mathbf{y}_{\text{now}})$. The pair is sufficient only in the normal case, as in this chapter's example; Section 6.4.2 showed the two drifting apart at a stated size elsewhere.

### 7.1.2 The same rows on the past and on the future

The rows $F$ and $h$ appear once in equation (7.1) and are read in two ways, split at the moment of deciding. Figure 7.1 draws this. Time runs from left to right, and the vertical line is now. The middle row of boxes is the state at successive steps, joined by the physics rows, drawn as small squares. Above them are the moves, each entering a physics row; below them are the

readings, and below those the fit and the policy. To the left of the line the readings have values, and the fit takes them. To the right nothing has been read, and the same boxes are drawn dashed.

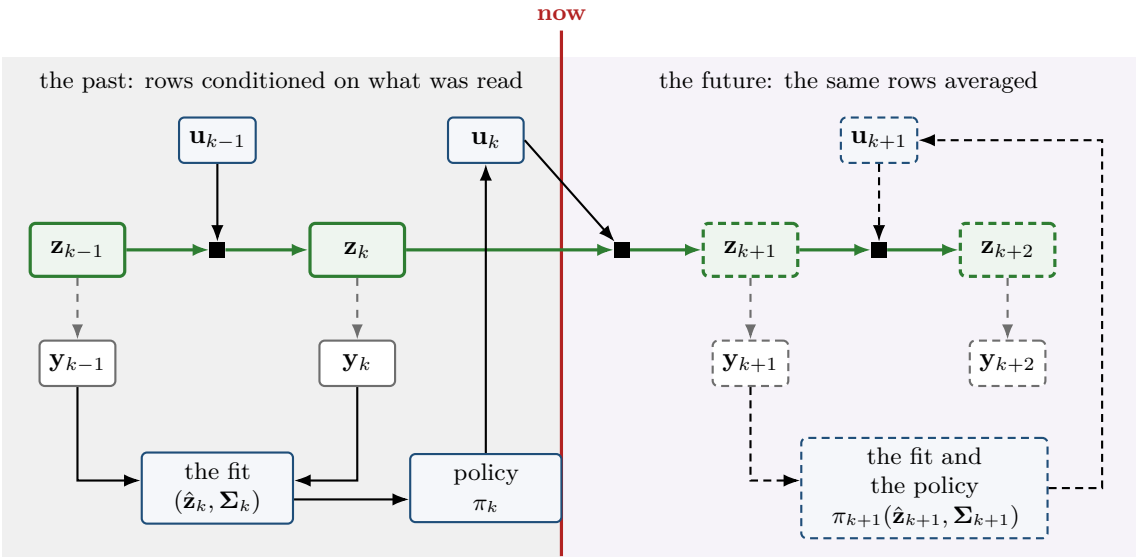



Figure 7.1: The problem (7.1) drawn across now. Green boxes: the state at successive steps. Black squares: the physics rows $F$. Blue boxes: the moves. Dashed arrows: the readings $h$. Left of the line the readings have values and the fit conditions on them. Right of the line the same rows hold quantities that have not happened, drawn dashed, and the expectation averages over them.

The figure states a rule the later chapters keep. On the past, the noise $\mathbf{w}_k$ is an unknown that the fit estimates: the rows $\mathbf{w}_k/\sigma_w$ belong in its objective, because a particular noise did happen and the readings say something about it. On the future, nothing has happened. The same symbol is a random quantity to be averaged over, with no value to estimate and no reason to come out favorably:

$$\begin{aligned} &\underbrace{\min_{\mathbf{z}_{0:k}} \ \tfrac{1}{2}\sum_{j<k} \frac{\|F(\mathbf{z}_{j+1}, \mathbf{z}_j, \mathbf{u}_j)\|^2}{\sigma_w^2} + \tfrac{1}{2}\sum_{j\le k} \frac{\|\mathbf{y}_j - h(\mathbf{z}_j, \mathbf{u}_{j-1})\|^2}{\sigma_\varepsilon^2}}_{\text{the past: conditioned}} \\ &\underbrace{\mathbb{E}_{\boldsymbol{\theta}|\mathbf{y}_{\text{now}}}\, \mathbb{E}_{\mathbf{w}_{k:}, \boldsymbol{\varepsilon}_{k+1:}} \Big[\sum_{j\ge k} \ell(\mathbf{z}_j, \mathbf{u}_j)\Big]}_{\text{the future: averaged}}, \end{aligned} \tag{7.2}$$

where $\boldsymbol{\theta}$ stands for whatever persists from the past into the future, such as an unknown gain. What persists is integrated against its posterior, not fixed at its estimate. For a model index, the most probable value on the past drops a mode, and the word is marginalized.

### 7.1.3 The four couplings

What separates equation (7.1) from a fit followed by a decision is that its lines act on each other. Figure 7.2 draws the same problem as a loop of four blocks: the policy, the graph, the readings and the fit. The black arrows are the order in which a step happens, clockwise. The numbered red arrows are the four couplings.

The first coupling says the move is not outside the model: it is an entry of the graph the fit reads. The second is what Chapter 5 met as a miss: the readings of a closed loop come from a

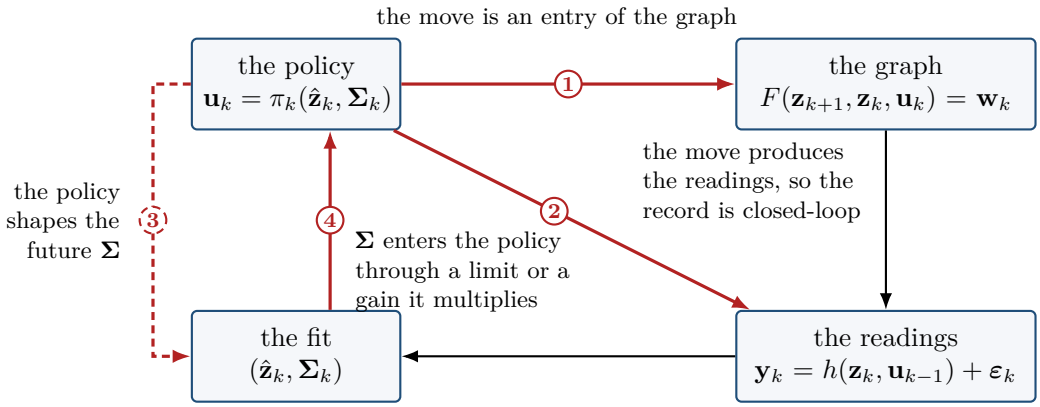



Figure 7.2: The four couplings of the problem (7.1). Black: the order of a step, from the graph to the readings to the fit. Red and numbered: 1, the move enters the physics rows; 2, the move reaches the readings through the plant; 3, dashed, the move changes the covariance the fit will have; 4, the covariance enters the next move through a limit or through a gain it multiplies.

plant the controller is moving, so the loop can hide a parameter or an offset. The third is the one a sheet of formulas never shows: what the plant reveals at the next step can depend on the move made now. The fourth is the band reaching the rule, by either of two routes. Through a limit it is Chapter 6's back-off carried into the rule: a wider band asks for a different move. Through a gain it multiplies it needs no limit and no learning: an uncertain effect of the move is a cost on moving. A controller drops some of these couplings, deliberately or not. The first two remain in every case below. When the third and fourth fall away the problem does come apart into a fit followed by a decision, and the example below shows the case in which that happens and two in which it does not. Optimism, also treated below, is not one of these cases: it is a planner's error.

### 7.1.4 The recursion

Without a limit, the problem (7.1) solves backward. The limit, not computability, is why the recursion is written without one: a joint limit is a single number across the readings, Section 7.3.3, and a recursion deals with one reading at a time. Let $V_k(\hat{\mathbf{z}}_k, \boldsymbol{\Sigma}_k)$ be the smallest expected consequence still to come from a belief at step $k$. The last move faces only its own consequence, and each earlier move faces its own plus whatever the next belief will be worth,

$$V_k(\hat{\mathbf{z}}_k, \boldsymbol{\Sigma}_k) = \min_{\mathbf{u}_k} \Big\{ \underbrace{\mathbb{E}\big[\ell(\mathbf{z}_k, \mathbf{u}_k) \mid \hat{\mathbf{z}}_k, \boldsymbol{\Sigma}_k\big]}_{\text{what the move costs now}} + \underbrace{\mathbb{E}_{\mathbf{y}_{k+1}}\big[V_{k+1}(\mathrm{Fit}(\hat{\mathbf{z}}_k, \boldsymbol{\Sigma}_k, \mathbf{u}_k, \mathbf{y}_{k+1}))\big]}_{\text{what the move leaves the future to know}} \Big\}. \tag{7.3}$$

This is the Bellman recursion on the belief. The second term averages over a reading not yet taken and asks what the belief after it would be worth, which is the preposterior of equation (4.73) inside a control problem. The move sits inside that term, so a move is judged partly as a measurement. For a state of any size, this recursion is not computable except in small cases, and that is why most of the rest of the book consists of restrictions. It is computable for the example that follows.

### 7.1.5 One loop and two moves (step 0)

The example is deliberately small, so that the general problem can be solved outright and its restrictions compared with it. **It is a one-node reading of the loop with stated choices, not a reduction derived from the closed-loop hall.** Chapters 12 to 15 use the full hall.

Here $x_k$ is the discharge air's deviation from its set point in K, and $u_k$ the valve's deviation from its operating position:

$$x_{k+1} = a\,x_k - \theta\,u_k + w_k, \qquad y_k = x_k + v_k, \tag{7.4}$$

with $\mathrm{sd}(w_k) = \mathrm{sd}(v_k) = 0.10\,\mathrm{K}$ and $k = 0, 1, 2$. The plant has two moves $u_0, u_1$ and one reading between them, $y_1$. The free entries are $x_0$, $x_1$, $x_2$ and the two moves; the coefficients are pinned, until Section 7.3.4 frees $\theta$. A policy leaves the second move free to respond, which is the closed-loop reading of a what-if, and a fixed sequence pins both moves in advance, which is the open-loop one.

**The choices.** One step is 60 s, taken equal to the valve actuator's time constant, so that $a = \mathrm{e}^{-1} = 0.36788$. The process gain is Chapter 5's fit, $K = 7.0996\,\mathrm{K}$ per unit of valve, which gives $\theta = (1 - a)K = 4.48779\,\mathrm{K}$ per unit; the minus sign says that opening cools. A single first-order node buys a problem small enough to solve exactly. It fails as a description of any particular hall, since it has no hall air, no integrator and no stop. The noises are stated, not fitted. So is the belief at the start, $x_0 \sim \mathcal{N}(1,\ 0.20^2)$. So are the costs, as spreads: being at the set point to 0.5 K and moving the valve by about 0.2 are worth the same,

$$\sum_k \ell = q\,(x_1^2 + x_2^2) + \rho\,(u_0^2 + u_1^2), \qquad q = 1/0.5^2 = 4.00, \quad \rho = 1/0.2^2 = 25.00. \tag{7.5}$$

A different pair of spreads gives different percentages throughout, which is one reason none of them is a property of control.

### 7.1.6 The control engineer's questions (step 1)

Table 7.1 lists the six questions. The first column is the question in the engineer's words, the second the general form it takes, the third its mathematical form on the example, the fourth its category, the fifth the computation that answers it and the last the subsection. The first question has been answered by the statement of the problem in Section 7.1.1, on which the others depend; the other five are the problem read under different conditions.

Table 7.1: The questions asked of the problem and of its example. The last row has no form of its own: it reads the others.

| Control engineer's question | General form | Mathematical form | Category | Computation | Answered in |
|---|---|---|---|---|---|
| What is the problem every controller is solving part of? | expected consequence over policies, (2.23) | the problem (7.1) | design | expected consequence | 7.1.1 |
| With no limits, does the band change the move? | the same, restricted, (7.3) | $u_k = K_k\hat{x}_k$, against $s_0$ | design | expected consequence | 7.3.1 |
| Can the future be fit like the past? | MAP against expectation, (7.2) | $u_0$ with the future noise left free | estimation | posterior | 7.3.2 |
| What does a response to the next reading save, under a limit? | chance constraint over policies, (6.1) | fixed, linear, free, held after every reading | design | posterior | 7.3.3 |
| When the plant is unknown, what does the band do to the move? | the recursion, (7.3) | certainty equivalent, cautious, exact | design | preposterior | 7.3.4 |
| Which later chapter is which restriction? | – | the map | – | all five | 7.3.5 |

### 7.1.7 The evidence and the costs as rows (step 2)

**The rows of the belief.** Three rows act on the entries $x_0$ and $x_1$: the belief at the start, the physics row of the first step with its noise as the tolerance, and the reading $y_1$. Each reading is taken as an independent normal given the state, which is what makes the fit exact here, and it fails for a sensor with a slow error of the kind Chapter 6 freed. With $*$ for an entry that can be nonzero, the pattern in the columns $x_0, x_1$ is

$$\mathbf{J} = \begin{bmatrix} * & 0 \\ * & * \\ 0 & * \end{bmatrix} \quad \begin{matrix} \text{belief at the start} \\ \text{physics of the first step} \\ \text{reading } y_1 \end{matrix} \tag{7.6}$$

and the move does not appear in it: $\theta u_0$ sits in the physics row's residual and has no line in the Jacobian. That zero is load-bearing. It says the moves cannot change this fit's covariance.

**The row of the unknown gain.** When $\theta$ is freed in Section 7.3.4, it takes a normal prior $\theta \sim \mathcal{N}(\hat{\theta},\ s^2)$ with $\hat{\theta} = 4.48779\,\mathrm{K}$ per unit and $s/\hat{\theta}$ set over a range, and $x$ is read exactly, so that the only entry is $\theta$. The physics row of the first step then reads $(x_1 - ax_0 + \theta u_0)/\sigma_w$, whose slope in $\theta$ is $u_0/\sigma_w$. The move is now in the line, not in the residual, and the second term of the recursion comes alive. Reading $x$ exactly buys a quadrature on one reading; a sensor's noise on $x$ would add a second unknown that the example does not need.

**The limit.** Section 7.3.3 asks for $x_2 \le 0.10\,\mathrm{K}$ with probability 0.95 and drops the $q$ terms, so that its cost is $\rho(u_0^2 + u_1^2)$, the valve travel needed to hold the limit.

## 7.2 The posterior of the state (step 3)

Every answer below uses one of two beliefs. They differ in whether the move has a line in the Jacobian, and all that follows turns on that.

### 7.2.1 The belief after a reading

The state is $\mathbf{z} = (x_0, x_1)^\mathsf{T}$ and the rows of the pattern (7.6), each divided by its standard deviation, are

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (x_0 - 1)/0.20 \\ (x_1 - ax_0 + \theta u_0)/0.10 \\ (x_1 - y_1)/0.10 \end{bmatrix}, \qquad \mathbf{J} = \begin{bmatrix} 5 & 0 \\ -3.67879 & 10 \\ 0 & 10 \end{bmatrix}, \tag{7.7}$$

in $\mathrm{K}^{-1}$, with $\mathcal{L} = \frac{1}{2}\mathbf{r}^\mathsf{T}\mathbf{r}$. The rows are linear, so one step of the normal equations $(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$ reaches the minimum, and the posterior is exactly normal. The information matrix of these normalized rows is $\mathbf{J}^\mathsf{T}\mathbf{J}$; with unnormalized rows it is $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$. Numerically,

$$\mathbf{J}^\mathsf{T}\mathbf{J} = \begin{bmatrix} 38.53353 & -36.78794 \\ -36.78794 & 200 \end{bmatrix}, \qquad \boldsymbol{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1} = \begin{bmatrix} 0.03148 & 0.00579 \\ 0.00579 & 0.00607 \end{bmatrix}. \tag{7.8}$$

Neither matrix contains $u_0$ or $y_1$. The mean $\hat{\mathbf{z}}$ depends on both, but the covariance is known before the reading is taken and whatever move is made.

The entry the second move needs is the last one, the variance of $x_1$ after the reading, $P = \Sigma_{22} = 0.00607\,\mathrm{K}^2$. The same number comes in the form Chapter 9 derives properly, used here and cited forward: the predicted variance $P^- = a^2 s_0^2 + \sigma_w^2 = 0.01541\,\mathrm{K}^2$, the gain $P^-/(P^- + \sigma_v^2) = 0.60651$ and the posterior $P = (1 - 0.60651)P^-$. The fit itself moves with the

reading: $\hat{x}_1 = m_1 + 0.60651\,(y_1 - m_1)$, with $m_1 = a\hat{x}_0 - \theta u_0$. Seen from now, before $y_1$ exists, $\hat{x}_1$ is a normal quantity with mean $m_1$ and variance

$$\mathrm{Var}(\hat{x}_1) = P^- - P = 0.00935\,\mathrm{K}^2, \qquad \mathrm{sd}(\hat{x}_1) = 0.09669\,\mathrm{K}, \tag{7.9}$$

the part of the prediction's variance that the reading will resolve. This is the only preposterior the first three answers need, and no move changes it.

### 7.2.2 The belief of an unknown gain

With $\theta$ free and $x$ read exactly, the state has one entry. The prior's slope and curvature, and the first step's row of Section 7.1.7, add to

$$\Lambda = \frac{1}{s^2} + \frac{u_0^2}{\sigma_w^2}, \qquad s_1^2 = \frac{1}{\Lambda}, \qquad \hat{\theta}_1 = s_1^2 \left( \frac{\hat{\theta}}{s^2} - \frac{u_0\,(x_1 - a x_0)}{\sigma_w^2} \right), \tag{7.10}$$

the exact normal posterior, since the row is linear in $\theta$. Here the move is in the information itself. A larger $|u_0|$ leaves a smaller $s_1^2$; $u_0 = 0$ teaches nothing. At $s/\hat{\theta} = 1.50$ the prior's spread is 6.7317 K per unit, and a first move of 0.02418 takes it to 3.5236. This is the third coupling, as an entry in a formula.

## 7.3 The answers (step 4)

The first question was answered by the statement of the problem. The four that follow take the example through it, in order of what they add: no limit and a known plant; the same with the future left free; a limit; an unknown gain. Then the map.

### 7.3.1 Separation, the case in which the band does not move the valve

**Control engineer's question.** With no limits, does the band change the move?

**General form.** The problem (7.1) with its last line removed, solved by the recursion (7.3), whose every step is an expected consequence (2.23) with the move as the action.

**Specialized form.** The recursion (7.3) on the example: $\ell$ is equation (7.5), the rows are (7.4), the belief is $(\hat{x}_k, P_k)$ of Section 7.2.1, and each step is the expected consequence (2.23). A known $\theta$ and a quadratic cost are the assumptions that enter here, and the last answer of this section shows what follows when the first is dropped.

Start at the last move, with $x_1 \sim \mathcal{N}(\hat{x}_1, P)$ after the reading. Since $x_2 = a x_1 - \theta u_1 + w_1$, and the mean of a square is the square of the mean plus the variance,

$$\mathbb{E}[q x_2^2 \mid \hat{x}_1, P] = q\,(a\hat{x}_1 - \theta u_1)^2 + q\,(a^2 P + \sigma_w^2). \tag{7.11}$$

The move appears only in the first term, and the band only in the second. The derivative of $\rho u_1^2$ plus equation (7.11) with respect to $u_1$ is $2\rho u_1 - 2q\theta(a\hat{x}_1 - \theta u_1)$. Setting it to zero,

$$u_1 = K_1 \hat{x}_1, \qquad K_1 = \frac{qa\theta}{\rho + q\theta^2}. \tag{7.12}$$

Putting this $u_1$ back gives what the last move leaves, with $q\hat{x}_1^2 + qP$ from $q x_1^2$ included,

$$V_1(\hat{x}_1, P) = S_1 \hat{x}_1^2 + c(P), \qquad S_1 = q + a^2 q - \frac{(qa\theta)^2}{\rho + q\theta^2}, \tag{7.13}$$

with $c(P) = q\,(P + a^2P + \sigma_w^2)$ a constant in everything that a move can reach. One step back, $\hat{x}_1$ has mean $m_1 = a\hat{x}_0 - \theta u_0$ and the variance of equation (7.9), neither of which involves the band in a way $u_0$ can touch, so

$$\rho u_0^2 + \mathbb{E}[V_1] = \rho u_0^2 + S_1\,(a\hat{x}_0 - \theta u_0)^2 + S_1\,\mathrm{Var}(\hat{x}_1) + c(P). \tag{7.14}$$

This has the shape of the last step with $S_1$ in place of $q$, and the same derivative gives

$$u_0 = K_0\hat{x}_0, \qquad K_0 = \frac{S_1 a\theta}{\rho + S_1\theta^2}. \tag{7.15}$$

This is the Riccati recursion, the backward pass that carries the coefficient of the quadratic cost-to-go from one step to the step before, here over two moves; Chapter 12 writes it for any state and horizon. The law is derived here: the gain is what the rows and the costs give, and nothing in it was tuned. In equation (7.15) neither $s_0$ nor $P$ appears. The band adds a constant to the cost, and a constant has no derivative.

**Building blocks.** $\hat{x}_0$ and $s_0$; $P$ and $\mathrm{Var}(\hat{x}_1)$ from equations (7.8) and (7.9); $a$, $\theta$, $q$, $\rho$.

**The plant's numbers.** With $qa\theta = 6.60386$ and $\rho + q\theta^2 = 105.56091$, equation (7.12) gives $K_1 = 0.062560$ per K. Then $S_1 = 4 + 0.54134 - 0.41314 = 4.12821$, so $S_1 a\theta = 6.81552$, $\rho + S_1\theta^2 = 108.14301$ and $K_0 = 0.063023$. At $\hat{x}_0 = 1\,\mathrm{K}$ the first move is $u_0 = 0.063023$.

The band of $x_0$ does not move it. A direct minimization of the expected cost over $u_0$ returns 0.063023, and across $s_0 = 0.05$, 0.20, 0.50 and 1.00 K the move stays there while the cost goes through 0.21393, 0.23529, 0.35347 and 0.77300. This is the one case in the chapter where the decision really is added on to the fit, and the engineer's reading of the opening scene is right in it. It holds here with linear rows, a known gain, a quadratic cost and no limit. These are sufficient on this example, and none is shown necessary: a limit that does not bind, for one, also leaves the move alone.

Recourse is still worth something here. The expected cost of the law is 0.23529. The best fixed sequence, $u_0 = 0.06302$ and $u_1 = 0.00532$, costs 0.23915, which is 1.64 % more on this example's noises and costs. The first move is the same and the second cannot respond. In closed form the gap is the coefficient of the last term of $S_1$ in equation (7.13) times the variance the reading resolves: $0.41314 \times 0.00935 = 0.00386$.

### 7.3.2 Optimism, the future fit like the past

**Control engineer's question.** Can the future be fit like the past?

**General form.** The two parts of equation (7.2) treat the noise differently. A planner built by reusing the fit on future rows makes this error: it minimizes over the noise. For a cost $J(\mathbf{u}, \mathbf{w})$ and noise with density $p(\mathbf{w})$, the two candidates are

$$\min_{\mathbf{u}}\ \mathbb{E}_{\mathbf{w}}\big[J(\mathbf{u}, \mathbf{w})\big] \qquad \text{against} \qquad \min_{\mathbf{u}}\ \min_{\mathbf{w}}\ \big[\tfrac{1}{2}J(\mathbf{u}, \mathbf{w}) - \ln p(\mathbf{w})\big]. \tag{7.16}$$

**Specialized form.** Equation (7.16) with $J$ the cost (7.5) and $w_k \sim \mathcal{N}(0, \sigma^2)$. First the case with no noise in the future at all. The state is $\mathbf{z} = (x_1, x_2, u_0, u_1)^{\mathsf{T}}$, the moves being entries like any other, and the rows are the physics at a tolerance $\sigma_p$ and the cost spreads of Section 7.1.5:

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (x_1 - a\hat{x}_0 + \theta u_0)/\sigma_p \\ (x_2 - ax_1 + \theta u_1)/\sigma_p \\ x_1/0.5 \\ x_2/0.5 \\ u_0/0.2 \\ u_1/0.2 \end{bmatrix}, \qquad \mathbf{J} = \begin{bmatrix} 1/\sigma_p & 0 & \theta/\sigma_p & 0 \\ -a/\sigma_p & 1/\sigma_p & 0 & \theta/\sigma_p \\ 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}. \tag{7.17}$$

Then $\mathcal{L} = \frac{1}{2}\mathbf{r}^{\mathsf{T}}\mathbf{r}$ is half the cost (7.5) on the physics, and the normal equations solve it in one step. This is a least-squares fit of a plan. It reproduces the law, because with no noise the recursion of Section 7.3.1 is a sequence of the same derivatives. So the form is shared: a linear-quadratic law is what a stack of rows gives. **This identity holds with no limit and no noise in the future. It is not the general solution**, and the next paragraph and the next two subsections show three ways of leaving it.

Now leave the noise in as entries, $w_0$ and $w_1$, each with its row $w_k/\sigma$, and the physics rows carrying $-w_k$. Treating them the way a fit treats the past takes the second form of equation (7.16). At one step, with cost-to-go weight $S$ and $m$ the mean the move leaves,

$$\min_{w}\Big[S\,(m+w)^2+\frac{w^2}{\sigma^2}\Big]=\frac{S}{1+S\sigma^2}\,m^2,\quad w=-\frac{S\sigma^2}{1+S\sigma^2}\,m;\qquad \mathbb{E}_w\big[S\,(m+w)^2\big]=S\,m^2+S\sigma^2. \tag{7.18}$$

Averaging adds a constant, which leaves the move alone. Minimizing scales the weight down to $S/(1+S\sigma^2)$, which does not. The planner has found noise that does part of the valve's work, at a cost much smaller than the valve's, and has planned on receiving it. The move is therefore the law's with $q$ and $S_1$ replaced by their shrunk values at each step.

**Building blocks.** The rows (7.17), $S_1$ and $K_0$ of Section 7.3.1, the shrunk weight of equation (7.18), and the expected cost of Section 7.3.1 to score the result.

**The plant's numbers.** With noise absent, the stacked rows give $u_0 = 0.063023$ and $u_1 = 0.005320$, the first equal to the law's at $\hat{x}_0 = 1$, and the second equal to $K_1$ times the mean of $x_1$ under the law, 0.08504 K. Table 7.2 leaves the noise free. Each column is a noise spread. The rows give the shrunk weight at the last step, the first move as a share of the law, the noise the minimization plans in K, and the excess expected cost of that first move with the law playing the second move, in two readings. In the first the plant's real noise stays at the example's 0.10 K whatever the planner assumes, so only the planner is wrong about the size. In the second the real noise equals the column's, so the planner has the size right and errs only by minimizing over it. Figure 7.3 draws the second and third rows against the noise spread on a logarithmic axis, with the law's ratio of one and a planned noise of zero as dotted references.

Table 7.2: The first move with the future noise left free, equation (7.16), second form. These are this example's values.

| Noise sd [K] | 0.01 | 0.05 | 0.10 | 0.20 | 0.50 |
|---|---|---|---|---|---|
| $q/(1+q\sigma^2)$ [K$^{-2}$] | 3.99840 | 3.96040 | 3.84615 | 3.44828 | 2.00000 |
| $u_0$ / the law [–] | 0.9999 | 0.9976 | 0.9905 | 0.9630 | 0.8065 |
| Planned $w_0$ [K] | −0.00004 | −0.00088 | −0.00348 | −0.01352 | −0.07078 |
| Excess, real sd 0.10 [%] | 0.0000 | 0.0010 | 0.0165 | 0.2501 | 6.8372 |
| Excess, real sd as column [%] | 0.0000 | 0.0014 | 0.0165 | 0.1226 | 0.7354 |

The optimizer plans on noise the plant has not promised, and the larger the noise the further it leans. The size here is about 1 % of the move at the example's noise and 19 % at five times that noise. At the example's noise it costs 0.0165 % of the expected cost, because a cost is flat near its minimum, and when the real noise is as large as the planner's it costs 0.12 % at 0.20 K and 0.74 % at 0.50 K. These are this example's sizes and not a general verdict that the error is small. The error is structural and its cost depends on the plant. On this plant more moves do not enlarge it, since $a = 0.36788$ forgets quickly: the ratio stays at 0.9905 for 2, 5 and 40 moves at the example's noise. On a slower plant with the same $K$ and $a = 0.95$ the same recursion gives 0.9174, 0.7815 and 0.6342 over those horizons, so there it grows with the

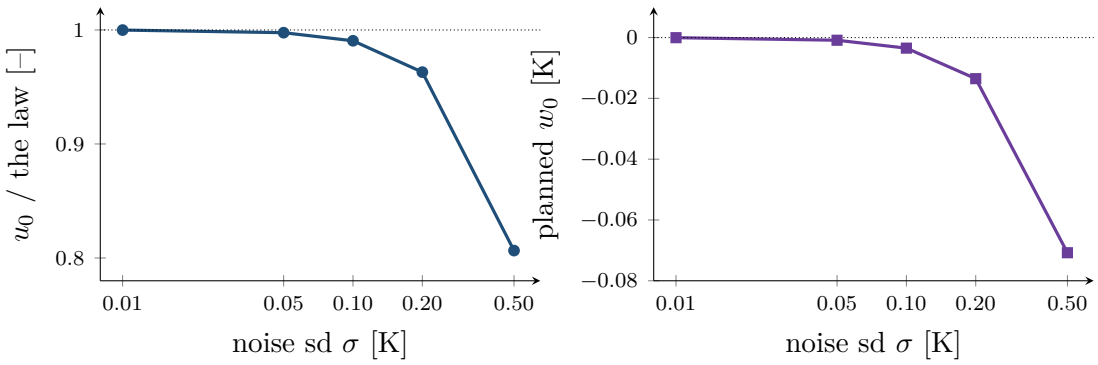



Figure 7.3: The optimism of Table 7.2. Left: the first move as a share of the law. Right: the first-step noise the minimization plans on, negative because a cooling noise helps. Dotted: the law's ratio of one, and no planned noise.

horizon. A planner that keeps the physics as soft rows over the future has this error built in, whatever the fit's quality on the past.

### 7.3.3 Recourse under a limit

**Control engineer's question.** What does a response to the next reading save, under a limit?

**General form.** The chance constraint (6.1), with $f$ a quantity produced by a policy,

$$\min_{\pi} \ \mathbb{E}\Big[\sum_{k} \ell(\mathbf{u}_k)\Big] \quad \text{subject to} \quad \Pr\big(f(\mathbf{z};\pi) \le f^* \mid \mathbf{y}_{\text{now}}\big) \ge 1-\alpha. \tag{7.19}$$

**Specialized form.** Equation (7.19) with $f = x_2$, $f^* = 0.10\,\text{K}$, $\alpha = 0.05$ and $\ell = \rho u^2$, on the rows (7.4). The probability is taken at now, over everything including the reading, so it is the equation (6.1) of step 0. The reading then splits it. Given $y_1$, the state $x_1$ is $\mathcal{N}(\hat{x}_1, P)$ by Section 7.2.1, so $x_2 = ax_1 - \theta u_1 + w_1$ is normal too,

$$x_2 \mid y_1 \sim \mathcal{N}\big(a\hat{x}_1 - \theta u_1,\ a^2P + \sigma_w^2\big), \qquad v(y_1) = 1 - \Phi\left(\frac{0.10 - a\hat{x}_1 + \theta u_1}{\sqrt{a^2P + \sigma_w^2}}\right), \tag{7.20}$$

where $v$ is the conditional violation after that reading. By the total probability rule the overall violation is its average over the readings,

$$\Pr(x_2 > 0.10) = \mathbb{E}_{y_1}\big[v(y_1)\big]. \tag{7.21}$$

Equation (7.21) is why this limit does not split across readings the way a cost does. A cost at the second step is a sum over steps, and the recursion deals with each reading on its own. The overall limit is a single number across all the readings, so risk spent after one reading is risk not available after another, and the plan cannot be made one reading at a time. The example limits $x_2$ alone, so a joint limit over time is not exercised here; Chapter 18 owes it. Four policies read this.

*The fixed sequence.* No reading is used, so $x_2$ has the mean $a^2\hat{x}_0 - \mathbf{g}^\mathsf{T}\mathbf{u}$ with $\mathbf{g} = (a\theta,\ \theta)^\mathsf{T}$ and the variance $a^2P^- + \sigma_w^2$. By the back-off (6.3) the constraint is $\mathbf{g}^\mathsf{T}\mathbf{u} \ge a^2\hat{x}_0 + 1.6449\ \text{sd}(x_2) - 0.10$, and the smallest $\mathbf{u}$ meeting it lies along $\mathbf{g}$.

*The second move linear in the reading.* Let $u_1 = c + \kappa\,(\hat{x}_1 - m_1)$. Then $x_2$ is still normal overall, with mean $am_1 - \theta c$ and variance $(a - \theta\kappa)^2\,\text{Var}(\hat{x}_1) + a^2P + \sigma_w^2$, and the three numbers $u_0$, $c$, $\kappa$ minimize $\rho(u_0^2 + c^2 + \kappa^2\,\text{Var}(\hat{x}_1))$ under the back-off, mean plus 1.6449 spreads at most

0.10 K. The constraint binds. With a multiplier on it, setting the three derivatives to zero gives two facts. The derivatives in $u_0$ and $c$ differ only by the factor $a$ in the mean, so $u_0 = ac$. The derivative in $\kappa$ gives $\kappa = 1.6449\,c\,(a - \theta\kappa)/\,\mathrm{sd}(x_2)$: the slope is the reduction of the spread bought per unit of valve travel. What is left is one equation for $c$, the binding constraint.

*The second move free after each reading.* This is the problem itself. One number, equation (7.21), ties all the readings together, and a multiplier unties them. A multiplier $\lambda$ is a price: add $\lambda\,v$ to the cost, so that risk can be bought at a stated rate instead of being limited. With a price, each reading has its own small problem, minimize $\rho u_1^2 + \lambda\,v$, which needs nothing from the other readings. A higher price buys less risk, so $\lambda$ is raised until the average violation is 0.05, and then $u_0$ is chosen. The method is stated in Section 7.4.2.

*The limit held after every reading.* This one asks for more: $v(y_1) \le 0.05$ at every reading, not on average. Each reading then has its own back-off, and $u_1 = (a\hat{x}_1 + 1.6449\ \mathrm{sd}(x_2 \mid y_1) - 0.10)/\theta$, which is the linear law with the particular slope $\kappa = a/\theta$. It is a stronger guarantee, and its cost is not comparable as a saving.

**Building blocks.** $P$ and $\mathrm{Var}(\hat{x}_1)$ of equations (7.8) and (7.9); the two spreads of $x_2$; the quantile $\Phi^{-1}(0.95) = 1.6449$.

**The plant's numbers.** The spread of $x_2$ is 0.10994 K with no reading used and 0.10402 K after the reading, so the back-offs are 0.18083 and 0.17110 K. For the fixed sequence, $\mathbf{g} = (1.65096,\ 4.48779)^{\mathsf{T}}$, $\mathbf{g}^{\mathsf{T}}\mathbf{g} = 22.86591$ and $\mathbf{g}^{\mathsf{T}}\mathbf{u}$ must reach $0.13534 + 0.18083 - 0.10 = 0.21616$. Table 7.3 gives the four policies and, last, the fixed sequence re-solved, which is taken up below. Each row is a policy; the columns give the guarantee it provides, its first move, its second move at the mean reading, and its cost. Figure 7.4 draws them. The upper panel has one bar per policy, its length the cost, on an axis starting at zero, with the stronger guarantee in a lighter fill. The lower panel has the reading on the horizontal axis, in standard deviations of $\hat{x}_1$, and on the vertical axis the conditional violation (7.20) under the fixed sequence, the linear law and the free move. The horizontal line is 0.05, which is also where the fourth policy lies at every reading.

Table 7.3: Policies for $x_2 \le 0.10$ K at 0.95. The first three hold the same guarantee and are comparable; the last two hold a stronger one. These are this example's costs.

| Policy | Guarantee | $u_0$ [–] | $u_1$ at mean [–] | Cost [–] |
|---|---|---|---|---|
| Fixed sequence | once, overall | 0.01561 | 0.04243 | 0.05109 |
| $u_1$ linear in the reading | once, overall | 0.01495 | 0.04065 | 0.04776 |
| $u_1$ free after each reading (the problem itself) | once, overall | 0.01495 | 0.04058 | 0.04776 |
| Held after every reading | conditional, stronger | 0.01491 | 0.04052 | 0.04816 |
| Fixed sequence, re-solved after the reading | conditional, stronger | 0.01561 | 0.04026 | 0.04818 |

*What the response saves.* Against the fixed sequence carried out whatever is read, the linear law saves 6.52 %. The linear law and the free move agree to the five decimals printed, 0.04776, and the free cost lies 0.00000096 below the linear one, Section 7.4.2, so on this example the linear law gets nearly all of the saving. The saving grows with what the reading resolves. Table 7.4 repeats the comparison for three values of $s_0$. Each row is a value; the columns give the spread of $\hat{x}_1$, the first move of each, the two costs and the saving. The 6.52, 19.36 and 41.37 % are this example's, and they depend on $s_0$ directly.

The third and fourth columns of that table are the fourth coupling, in numbers. With no limit, the first move stood at 0.063023 whatever $s_0$ was. Under a limit it moves with the band:

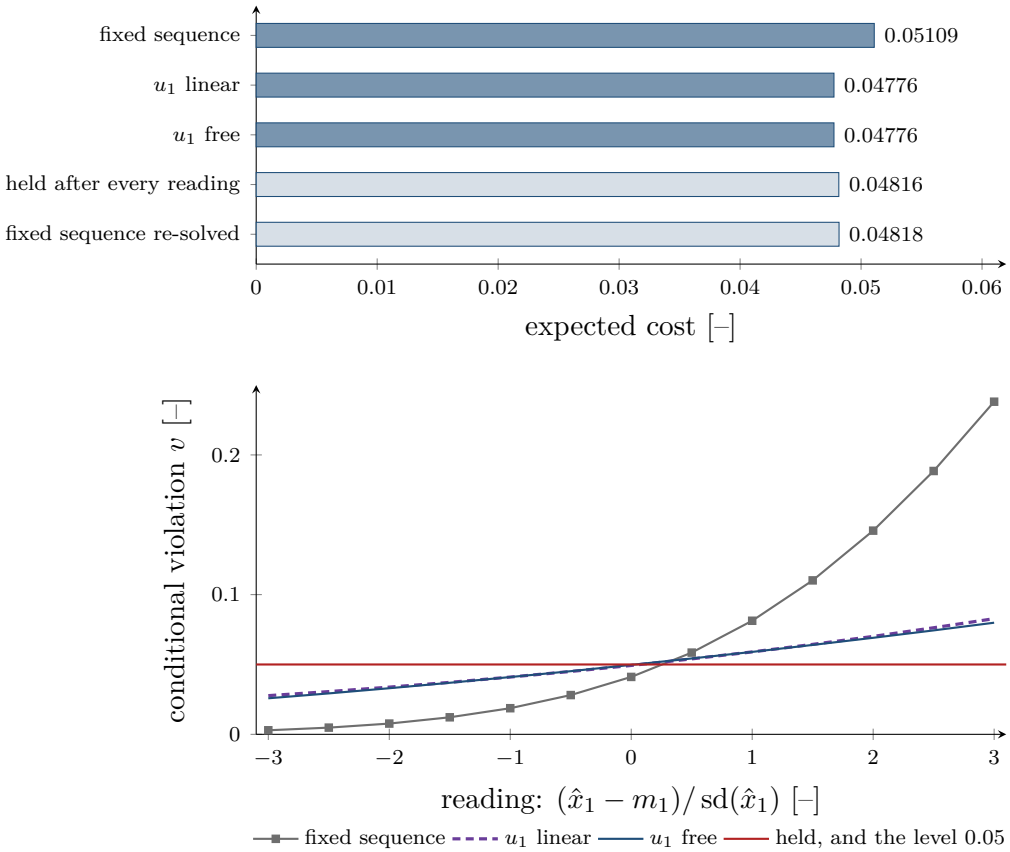



Figure 7.4: Upper panel: the costs of Table 7.3, with the stronger guarantees in the lighter fill. Lower panel: the conditional violation (7.20) after each reading. Each of the three lower curves averages to 0.05 over the readings by equation (7.21). The curve of the fourth policy, and of the fixed sequence re-solved, is the horizontal line.

Table 7.4: The saving of the linear law over the fixed sequence as the band of $x_0$ widens, for the same limit. This example's values.

| $s_0$ [K] | $\mathrm{sd}(\hat{x}_1)$ [K] | $u_0$ fixed [–] | $u_0$ linear [–] | Cost fixed [–] | Cost linear [–] | Saving [%] |
|---|---|---|---|---|---|---|
| 0.2 | 0.09669 | 0.01561 | 0.01495 | 0.05109 | 0.04776 | 6.52 |
| 0.5 | 0.18892 | 0.01754 | 0.01525 | 0.06454 | 0.05204 | 19.36 |
| 1.0 | 0.36875 | 0.02301 | 0.01584 | 0.11102 | 0.06510 | 41.37 |

from 0.01561 to 0.02301 for the fixed sequence, and less for the linear law, which leaves more of the work to the second move. $\mathbf{\Sigma}$ has entered the policy.

*The overall limit is not a limit after each reading.* In the lower panel the free move's conditional violation is 0.0330, 0.0497 and 0.0691 at $-2$, 0 and $+2$ standard deviations, and 48.5 % of readings leave it above 0.05. The solution means it. A high reading needs a large move, and the cost of a move is quadratic, so it is cheaper to accept more risk there and buy it back at low readings, which is what a slope of 0.06081 rather than the full-cancellation $a/\theta = 0.08197$ does in the linear law. Once a reading is in, the plan made before it no longer holds the same guarantee at that reading. A controller given this plan and then asked after a high reading whether the limit still holds at 0.95 would have to say no. A controller that restates the limit after each reading has moved to the conditional form, which is the next paragraph's.

*The stronger form.* Holding the limit after every reading is the form a recursion takes naturally, since each reading then stands alone. It costs 0.04816 against 0.04776 for the overall form, 0.85 % more on this example, and for that it provides the conditional guarantee. Since it is the linear law at one particular slope, it had to cost at least as much. The two guarantees are different, so which one a plant needs is a decision about the limit and not a result of the computation.

*The fixed sequence, re-solved.* The 6.52 % prices a sequence carried out whatever is read, and that is not what a re-solving controller does. Keep the fixed sequence's first move, 0.01561, and after the reading solve again for the one move left, against the limit as it then stands. Then the conditional violation is 0.0500 after every reading, by the closed form, so it is 0.0500 overall, and the cost is 0.04818, not 0.05109. That is 0.88 % above the free move and 0.00002 above the held policy, which chose its first move knowing the reading would come. So a re-solved sequence keeps the recourse it realizes and gives up only the recourse it did not anticipate, and on this example, with one reading ahead, that is a small price. The 6.52 % does not apply to it.

### 7.3.4 The band and an unknown plant

**Control engineer's question.** When the plant is unknown, what does the band do to the move?

**General form.** The recursion (7.3) with both of its terms, in which the second term is the preposterior of the belief a move will leave, each step an expected consequence (2.23). Nothing about the fit is restricted.

**Specialized form.** The recursion (7.3) on the example, with the unknown $\theta$ of Section 7.2.2 as the state, the cost (7.5), $x$ read exactly, and each step the expected consequence (2.23). The assumption of a normal prior on $\theta$ enters here. It buys the closed update (7.10), and it fails at the wide end of the range, where the prior gives real weight to a gain of the wrong sign (Section 7.6.3). Everything new comes from one line. At the last move, with $\theta$ unknown,

$$\mathbb{E}\big[(ax_1 - \theta u_1 + w_1)^2\big] = (ax_1 - \hat{\theta}_1 u_1)^2 + s_1^2 u_1^2 + \sigma_w^2, \tag{7.22}$$

because $\mathbb{E}[(\theta u)^2] = u^2\mathbb{E}[\theta^2] = u^2(\hat{\theta}^2 + s^2)$. Compare equation (7.11). There the band stood alone, as a constant, because a band on $x$ is added to the move's effect. Here the band is multiplied by $u^2$, because the unknown multiplies the move. So the band is now a cost on moving, and separation is gone with no limit in sight. The same derivative as before gives

$$u_1 = K_1(s_1)\, x_1, \qquad K_1(s_1) = \frac{qa\hat{\theta}_1}{\rho + q(\hat{\theta}_1^2 + s_1^2)}, \tag{7.23}$$

which is equation (7.12) with the band in the denominator, and leaves, with $qx_1^2$ included as in equation (7.13),

$$V_1(x_1, \hat{\theta}_1, s_1^2) = qx_1^2 + qa^2x_1^2 + q\sigma_w^2 - \frac{(qax_1\hat{\theta}_1)^2}{\rho + q(\hat{\theta}_1^2 + s_1^2)}. \tag{7.24}$$

Three routes then choose the first move, and a fourth is kept as a measure.

*Certainty equivalent.* Both moves ignore the band and use the current estimate as if it were true: $u_0$ is the law's, and $u_1$ uses $\hat{\theta}_1$ with $s_1^2$ dropped. Its cost is the true expected cost of those moves.

*Cautious.* The first move takes the band into account but plans as if nothing will be learned, with $\hat{\theta}_1 = \hat{\theta}$ and $s_1 = s$ in equation (7.24). That makes $V_1$ quadratic in $x_1$ again, with the coefficient $S_1(s) = q + qa^2(1 - q\hat{\theta}^2/(\rho + q(\hat{\theta}^2 + s^2)))$, and by equation (7.22) one step earlier,

$$u_0 = \frac{S_1(s)\, a\hat{\theta}}{\rho + S_1(s)\,(\hat{\theta}^2 + s^2)}\, x_0. \tag{7.25}$$

After the reading, the second move uses the real $\hat{\theta}_1$ and $s_1^2$ through equation (7.23).

*No update.* The cautious first move, and a second move that reads $x_1$ but keeps the prior on $\theta$, equation (7.23) at $\hat{\theta}$ and $s^2$. Its cost is the true expected cost of those moves. This route is the measure of what the update is worth, since it is the only one of the four that does not use the reading to learn about $\theta$.

*Exact.* The first move minimizes the whole recursion, with $x_1 \sim \mathcal{N}(ax_0 - \hat{\theta}u_0,\ u_0^2s^2 + \sigma_w^2)$ and the update (7.10) inside the second term:

$$J(u_0) = \rho u_0^2 + \mathbb{E}_{x_1}\left[V_1\big(x_1, \hat{\theta}_1(u_0, x_1), s_1^2(u_0)\big)\right]. \tag{7.26}$$

Here $u_0$ acts on the second term through $s_1^2(u_0)$, which is the dual effect. There is no closed form, and the expectation is a Gauss–Hermite quadrature over the one reading.

**Building blocks.** The update (7.10); the moments (7.22); the cost-to-go (7.24); the quadrature, with the grid in Section 7.4.2.

**The plant's numbers.** At $s/\hat{\theta} = 1$ the band equals the gain, $s^2 = \hat{\theta}^2 = 20.14023$, and equation (7.23) at the prior's values gives 0.03548 against the known-gain 0.062560. Table 7.5 gives the routes. Each row is a width of the prior; the columns give the first move and the expected cost of each route, the cost with no update, and the exact route's gain over the cautious one. The last three rows take the same normal truncated to $\theta > 0$, whose mean and spread differ, so that the certainty-equivalent move is that prior's own. Figure 7.5 draws the three routes on the normal prior. Its horizontal axis is $s/\hat{\theta}$ in both panels; the upper panel is the first move, the lower the expected cost, and the three lines in each are the three routes.

*Caution is worth something on this example, and how much depends on the prior.* The certainty-equivalent move does not change with the width of the prior, and its cost climbs. At $s/\hat{\theta} = 0.50$, where the prior gives the wrong sign a probability of 0.0228, the exact route's first move is 15.64 % below it and the exact route pays 4.20 % less. At $s/\hat{\theta} = 1.50$ the move is 61.63 % below it and the exact route pays 47.27 % less; put the other way, the certainty-equivalent cost is 89.63 % above the exact one. That wide end rests on a probability of 0.2525 that the valve warms. On the same normal truncated to $\theta > 0$ the two figures at 1.50 are 28.04 %, from that prior's own certainty-equivalent move, and 15.91 %.

The comparison is fair to certainty equivalence. Its first move followed by the optimal second move costs 0.95604 against 0.95645 at the wide end, so its second move adds only 0.00041, and the gap is the first move's. Certainty equivalence fails here with no limit in the problem, because $\theta$ multiplies $u$. The two bands are different quantities. In Section 7.3.1 the band is on

Table 7.5: The first move and the expected cost with $\theta$ unknown, at $x_0 = 1\,\mathrm{K}$, on the normal prior and, below the rule, on the same normal truncated to $\theta > 0$. These are this example's values, over two moves.

| | Certainty equivalent | | Cautious | | Exact | | No update | Exact's |
|---|---|---|---|---|---|---|---|---|
| $s/\hat{\theta}$ [–] | $u_0$ [–] | Cost [–] | $u_0$ [–] | Cost [–] | $u_0$ [–] | Cost [–] | cost [–] | gain [%] |
| 0.10 | 0.06302 | 0.21389 | 0.06255 | 0.21387 | 0.06257 | 0.21387 | 0.21387 | 0.00 |
| 0.25 | 0.06302 | 0.23226 | 0.06019 | 0.23138 | 0.06027 | 0.23138 | 0.23146 | 0.00 |
| 0.50 | 0.06302 | 0.29961 | 0.05302 | 0.28704 | 0.05317 | 0.28704 | 0.28783 | 0.00 |
| 1.00 | 0.06302 | 0.55282 | 0.03582 | 0.41570 | 0.03637 | 0.41565 | 0.43417 | 0.01 |
| 1.50 | 0.06302 | 0.95645 | 0.02320 | 0.50466 | 0.02418 | 0.50437 | 0.54149 | 0.06 |
| 0.50, $\theta > 0$ | 0.06207 | 0.28073 | 0.05350 | 0.27135 | 0.05363 | 0.27135 | 0.27172 | 0.00 |
| 1.00, $\theta > 0$ | 0.05381 | 0.32429 | 0.04086 | 0.28942 | 0.04102 | 0.28941 | 0.29111 | 0.00 |
| 1.50, $\theta > 0$ | 0.04487 | 0.34205 | 0.03213 | 0.28762 | 0.03229 | 0.28761 | 0.29099 | 0.00 |

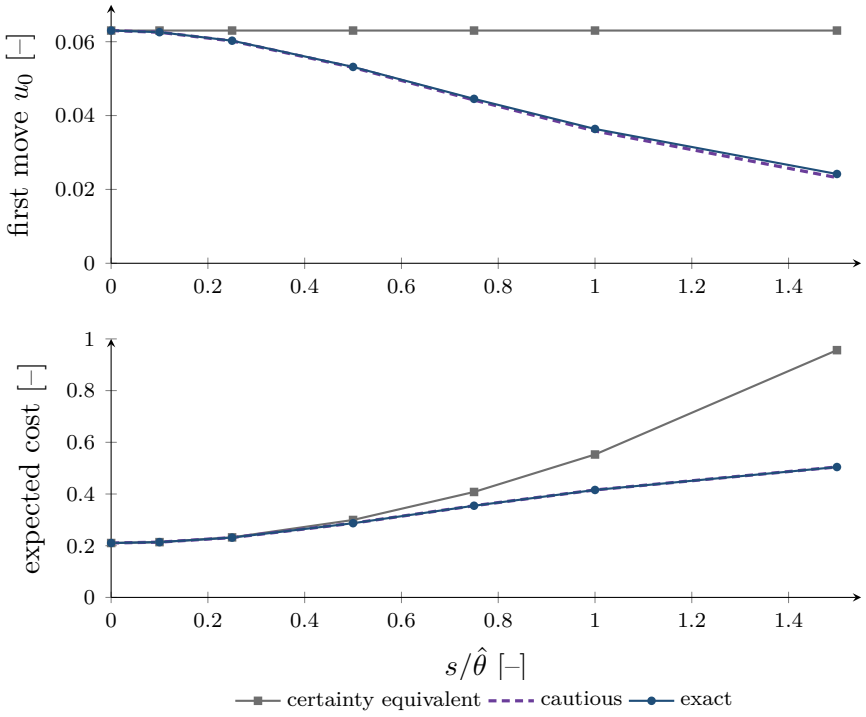



Figure 7.5: The three routes of Table 7.5, with the two further widths the table leaves out, against the width of the prior on $\theta$. Upper panel: the first move. Lower panel: the expected cost, each route's moves scored by the same expectation.

$x_0$, and it left the first move alone; here it is on $\theta$, with $x$ read exactly, and it cuts the move. At a narrow prior it hardly matters: at $s/\hat{\theta} = 0.10$ the exact route pays 0.01 % less.

*The lean of the first move was worth under* 0.1 %*; using the reading was not.* The exact route's first move does lean toward learning: at the wide end it is 0.02418 against the cautious 0.02320, and it leaves $\theta$'s spread at 3.5236 against 3.6296. That lean gains 0.06 % of the cost. But both of those routes use the reading at the second move, so this prices only the lean. With no update the cost at the wide end is 0.54149 against 0.50437, which is 7.36 % more, and choosing the best first move for that route recovers little, 0.54047 and 7.16 %. The update is worth 4.46 % at $s/\hat{\theta} = 1.00$ and 0.28 % at 0.50; on the prior truncated to $\theta > 0$ it is worth 1.18 % at 1.50. In the other direction, at $x_0 = 0$ with $s/\hat{\theta} = 0.50$, where the certainty-equivalent move is zero and teaches nothing, the exact first move is also 0.00000 with cost 0.08194, so no probing move is worth making there. With one move left after the reading, little remains to repay a probe, and a probe is paid for now.

*What this does and does not settle.* Two moves cannot show dual control. They are the shortest problem in which the term exists at all, and they give what is learned one move to pay off. The result says nothing against the dual effect in longer problems. Chapter 22 owes a horizon or a plant in which it pays, or a plain statement that it does not.

### 7.3.5 The map of restrictions

**Control engineer's question.** Which later chapter is which restriction?

Table 7.6 is the book's plan, written against equation (7.1). Each row names a part of the book, the second column says what is restricted, whether policy class, horizon, plant class or objective, the third lists which of the couplings of Section 7.1.3 remain, and the last names the chapters. Most rows are restrictions; the ones that are not say so in the second column. The four answers above have already met several rows on the example: the known-gain row is Section 7.3.1, the re-solved and held rows are in Section 7.3.3, the affine row is its linear law, and the row keeping the second term is Section 7.3.4.

Table 7.6: The restrictions of the problem (7.1) and the chapters that work under them.

| Row | What is restricted (policy class, horizon, plant class, objective) | Couplings kept (of 1–4) | Chapters |
|---|---|---|---|
| Fixed law, gains given | policy class: no objective produced it | 1, 2 | 5, 8 |
| One move, no future | horizon: one move | 1, 4 | 4, 6, 16, 17 |
| The Fit line itself | not a restriction of the decision: the line is studied alone, with no policy chosen | 2 | 9–11 |
| Known gain, quadratic, no limit | plant class and objective; that $\mathbf{\Sigma}$ then leaves the policy is a result, not an assumption | 1, 2 | 12–14 |
| The stop kept | plant class: as the row above, with the stop, which ends linearity | 1, 2, 4 | 15 |
| A fixed sequence, re-solved | policy class: anticipated recourse; the limit becomes conditional. Realized recourse is kept | 1, 2, 4 | 18, 19 |
| Limit held after every reading | the guarantee: conditional in place of joint | 1, 2, 4 | 18–20 |
| Affine in the disturbance | policy class: part of recourse | 1, 2, 4 | 20 |
| A re-solved plan with a store | the same policy class, on a plant with a tank or a battery in its state | 1, 2, 4 | 21 |
| The second term kept | no coupling is restricted: the future $\mathbf{\Sigma}$ is chosen; horizon and plant class stay small | 1–4 | 22–25 |
| The model index in the state | not a restriction: an enlargement of the state. Chapter 28's worst case is a changed objective | 1, 2, 4 | 26–28 |
| Levels tied by a reference map | policy class: global optimality across levels | each level's own | 29, 30 |

A controller's row is thus not a rival of the problem but a place on this table. The PI row of the opening scene is its first line. A model predictive controller is the row of the fixed sequence re-solved, and the offer of one should say so: a re-solved sequence keeps realized recourse, so its price on the example is not the 6.52 % of a sequence carried out whatever is read.

## 7.4 The checks (step 5)

The example has closed forms and quadrature at every point, so its tests compare routes and test budgets. There are no draws in this chapter and no truth row: the example has no record to hold a truth.

### 7.4.1 The stacked rows against the recursion

The stacked rows (7.17) at $\sigma_p = 10^{-7}$ give $u_0 = 0.063023$, the recursion's $K_0\hat{x}_0$, and $u_1 = 0.005320$, its $K_1 \times 0.08504$. Left free, the noise row $w_k/\sigma$ with the physics row carrying $-w_k$ is the same as a physics row at the tolerance $\sigma$, so Table 7.2 is also this book's limit check: as the tolerance falls to 0.01 K the ratio reaches 0.9999. The closed form of equation (7.18), applied at each step, reproduces the first moves of the stacked solve to the digits printed: 0.06302, 0.06287, 0.06242, 0.06069 and 0.05083. At the other end of Section 7.3.4, the prior at $s/\hat{\theta} = 0.0001$ returns all three routes to one move, 0.06302, and one cost, 0.21044, and the closed form (7.25) reproduces the cautious column at all seven widths. That cost lies below the 0.23529 of Section 7.3.1 because this section knows $x_0$ and reads $x_1$ exactly, and the gap is accounted for: removing the band of $x_0$ takes the law's cost from 0.23529 to 0.21250, and removing the reading's noise takes it to 0.21044.

### 7.4.2 A restriction cannot beat what it restricts

Each comparison in this chapter has an order that the mathematics fixes: the policy no worse than the fixed sequence, the free move no worse than the linear law and the linear law no worse than the fixed sequence, and the exact route no worse than the cautious or the certainty-equivalent. A value out of that order is a numerical error, whatever else it looks like.

The free move is computed without a grid on the move. For each reading the second move is found by Newton's method on $\rho u_1^2 + \lambda\, v$, starting from a coarse scan for the global minimum, because that function is not convex in $u_1$. The readings are Gauss–Hermite nodes. The multiplier is halved 60 times to set the average violation to 0.05, and $u_0$ is found by a bounded search. The cost is 0.04775828 at 101, 161 and 201 nodes. The free first move is 0.01495, its second move 0.04058 at the mean reading and 0.05874 at +3 standard deviations, where the linear law gives 0.05829. The two laws are not the same function, and 48.5 % of the readings leave the free move above 0.05. The linear law, by a constrained minimization, costs 0.0477592 and the free move 0.0477583, so the free cost is not above the linear one: it lies 0.00000096 below it, on the side the order requires. The first move of the exact route in Section 7.3.4 was searched over 3401 values of $u_0$ from $-0.5$ to 1.2 and then refined, with 161 Gauss–Hermite nodes, and its order held at all seven widths.

### 7.4.3 Budgets against closed-form moments

The policy's expected cost is the sum of its four terms: $q\mathbb{E}[x_1^2] = 0.09058$, $\rho u_0^2 = 0.09930$, $\rho\mathbb{E}[u_1^2] = 0.00162$ and $q\mathbb{E}[x_2^2] = 0.04379$, which sum to 0.23529. Its gap to the fixed sequence has the closed form 0.00386 of Section 7.3.1. The fixed sequence of Section 7.3.3 holds at 0.9500 by the normal tail. The linear law has mean $-0.07177$ K and spread 0.10443 K, so its mean plus 1.6449 spreads is 0.10000 K, the limit. The variance of $x_2$ splits by the total variance rule: $0.01209 = 0.01082 + 0.00127$, the last being $a^2\,\mathrm{Var}(\hat{x}_1)$. Equation (7.21) checks directly:

averaging the conditional violation over the readings by Gauss–Hermite gives 0.05000 for the fixed sequence, the linear law and the free move. The linear law's stationary conditions check too: $u_0/c = 0.36788 = a$, and $1.6449\,c\,(a - \theta\kappa)/\,\mathrm{sd}(x_2) = 0.06081$, its slope.

## 7.5 The report (step 6)

The sheets answer one question: the problem (7.1), of which each is a restriction. Separation is one case of it, shown here on linear rows with a known gain, a quadratic cost and no limit, and in that case the band, there on $x_0$, left the move alone. With an unknown gain and $x$ read exactly, the band, now on $\theta$, changed the move: at $s/\hat{\theta} = 0.50$, where the wrong sign has probability 0.0228, the exact first move is 15.64 % below the certainty-equivalent one and pays 4.20 % less. The wide end, 1.50, gives 61.63 % and 47.27 % on a prior with probability 0.2525 of the wrong sign, and 28.04 % and 15.91 % with that sign removed. Over two moves the lean of the first move was worth under 0.1 % at every width; using the reading to update $\theta$ was worth 0.28 % at 0.50 and 7.36 % at 1.50, or 1.18 % with the sign removed. All of these percentages are this example's, on a constructed one-node reading of the loop with stated noises and costs.

### 7.5.1 The answers and what each rests on

Table 7.7 gives the answers in the order of Table 7.1. Each row is a question; the second column gives the answer; the third what it rests on. A band on the state enters each answer as an input, so there are no posterior bands on these values: the band is what is being varied.

Table 7.7: The report on the example. Each percentage is this example's, for its stated noises, costs and two moves.

| Control engineer's question | Answer | Rests on |
|---|---|---|
| What is the problem every controller is solving part of? | Equation (7.1): a minimization over policies, with the fit as a constraint | the fit as an approximation of the belief |
| With no limits, does the band change the move? | No: $K_1 = 0.062560$, $K_0 = 0.063023$, at every $s_0$. The fixed sequence costs 1.64 % more | known gain, quadratic cost, no limit |
| Can the future be fit like the past? | Only with no noise, where the stacked rows give 0.063023. With it free, the move falls 1 % at the example's noise and 19 % at five times it | this plant's fast $a$; slower plants differ |
| What does a response to the next reading save, under a limit? | 6.52 % for a linear law against a sequence carried out whatever is read; the free move lies 0.00000096 below the linear law. Holding the limit after every reading, a stronger guarantee, costs 0.85 % more than the overall form. The sequence re-solved costs 0.04818, not 0.05109 | $s_0 = 0.20$ K; the saving is 41.37 % at 1.0 K |
| When the plant is unknown, what does the band do to the move? | Shrinks it: at $s/\hat{\theta} = 0.50$ the exact move is 15.64 % below the certainty-equivalent one and pays 4.20 % less. At 1.50, 61.63 % and 47.27 %, or 28.04 % and 15.91 % for a positive gain. The first move's lean gains under 0.1 %; updating $\theta$ from the reading is worth 0.28 % at 0.50 and 7.36 % at 1.50 | a normal prior, with probability 0.0228 and 0.2525 of the wrong sign; two moves, so it cannot show dual control; Chapter 22 owes a case |
| Which later chapter is which restriction? | Table 7.6 | – |

### 7.5.2 What a controller's sheet should declare

This chapter has no plant to log, so its entry is a requirement in place of a commissioning-log entry. A controller's sheet should declare which restriction of equation (7.1) it works under and which couplings it keeps. The vendor's PI row should say that its gains were given and that no objective produced them. A fixed-sequence planner should say that it plans without recourse, and whether it re-solves. A planner that minimizes over future noise should say so. A sheet carrying a limit should say whether the limit holds once overall or after every reading, because those are different guarantees at different costs. None of this needs a better controller. It lets the commissioning log state what was assumed away.

### 7.5.3 Common mistakes

Table 7.8 lists five mistakes this chapter's example makes easy. Each row names the mistake, its form on the example and the correction.

Table 7.8: Five mistakes in reading a controller against the problem.

| Mistake | On the example | Correction |
|---|---|---|
| Treating separation as a general result | 0.063023 at every $s_0$ with $\theta$ known; a cut of 15.64 % at $s/\hat{\theta} = 0.50$ with $\theta$ unknown | check that the gain is known and no limit binds before dropping $\mathbf{\Sigma}$ |
| Fitting the future | planned $w_0 = -0.07078\,\mathrm{K}$ at a noise of $0.50\,\mathrm{K}$ | keep the rows on the future as an expectation; a free noise is a plan to receive it |
| Reading a stacked least-squares law as the general solution | it matches the recursion only with no limit and no noise | use it for what it is, the deterministic case |
| Comparing unlike guarantees | 0.04816 against 0.04776 as if it were a loss of the same kind | the row held after every reading gives more; state the guarantee beside the cost |
| Taking an overall limit as held after each reading | 0.0691 after a reading at +2 standard deviations | equation (7.21): it is an average over the readings |

## 7.6 The next measurement (step 7)

This chapter has no new candidate sensor, because the measurement it studies is already in the problem. The reading in the recursion is the measurement.

### 7.6.1 The move as a measurement

The preposterior of the earlier chapters, equation (4.73), ranks a sensor by the band it leaves. The second term of equation (7.3) does the same for a move, and equation (7.10) is the case in which the move has a line in the Jacobian: at the wide end, the exact move leaves a spread of 3.5236 on $\theta$ and the cautious one 3.6296, from the prior's 6.7317. From here on step 7 asks what a move would teach as well as what a sensor would. Within this chapter's two moves, the lean of the first move was worth under 0.1 %, and that answer belongs to the horizon; using the reading to update $\theta$ was worth 0.28 % at $s/\hat{\theta} = 0.50$ and 7.36 % at 1.50.

### 7.6.2 What the analysis showed

The four sheets suggested that estimating, tuning and adding a margin are separate jobs, with the band left in the report. The analysis finds one problem under all of them, and the band

reaches the valve through the fourth coupling, by a limit or by a gain it multiplies. On a loop with a known gain and no limit, the sheets' reading was right: the band moved the cost and not the move. A limit brought the band into the first move and made the response to the next reading worth paying for. An unknown gain brought it in with no limit and no learning, and the wider the band, the more ignoring it cost. Couplings one and two stayed throughout; optimism, the remaining result, is a planner's error and not a coupling. What the engineer takes back is a question to put to each controller on offer: which couplings does it keep, and does this plant have the others.

### 7.6.3 Limits of the model

**The example's percentages.** The example is a one-node reading of the loop with stated noises and costs. Its percentages are this example's, not properties of control. A longer horizon, another cost or a slower plant gives others, as the slower plant in Section 7.3.2 does.

**Computability.** The general problem is not computable beyond small cases: the belief is the state of the recursion, so the recursion runs over a space the size of the state's mean and covariance together. The problem is the map, and the later chapters are what can be computed.

**The fit inside the problem.** The fit is itself an approximation, as Section 6.4.2 showed by its drift from normal. Nothing in this chapter tests that, because its posteriors are exact.

**The model set.** The problem assumes the model set is right. Chapter 5's miss is a warning about that, and no objective repairs it.

**A normal prior on a gain.** At the wide end the prior on $\theta$ gives the wrong sign real weight: the probability of $\theta < 0$ is 0.0228 at $s/\hat{\theta} = 0.50$, 0.1587 at 1.00 and 0.2525 at 1.50. The figures at that end partly reflect a valve that might warm: with the same normal truncated to $\theta > 0$, the cut at 1.50 falls from 61.63 % to 28.04 %, what the exact route saves from 47.27 % to 15.91 %, and what the update is worth from 7.36 % to 1.18 %. The direction of the result rests on equation (7.22), not on the sign, and the chapter's headline is at 0.50 for this reason.

**The constructed example.** The example is constructed. It has no record, and the two-move horizon, the single reading and the absence of a stop are all choices. Chapter 8 returns to the loop as a loop, and asks whether the row given by the vendor is stable under the band.

## 7.7 What to retain

**Definition 7.1** (The control problem)**.** The control problem is a minimization of expected consequence over policies on the belief, with the physics, the readings and the fit as constraints and the limits as one joint chance constraint, equation (7.1). Rows on the past are conditioned and rows on the future are averaged. Without a limit it solves backward by the recursion (7.3), whose second term is the preposterior of the belief the move will leave.

Three conclusions carry beyond the example, as statements of form and not of size. The band does not move the valve where a known gain, a quadratic cost and no limit all hold on linear rows, and each of those can be checked on the sheet; none is shown necessary, and a limit that does not bind also leaves the move alone. Minimizing over future noise is a different problem from averaging over it, and its error is structural. And an overall limit is an average across the readings, so a guarantee has to state when it is taken.

Table 7.9 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 7.9: The terms of the chapter and where each is defined. Units for the example.

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| The control problem | eq. (7.1) | – | one minimization over policies | 7.1.1 |
| Past and future rows | eq. (7.2) | – | conditioned; averaged | 7.1.2 |
| The recursion | eq. (7.3) | – | cost now, and what is left to know | 7.1.4 |
| Spread of the fit | $\mathrm{Var}(\hat{x}_1)$, eq. (7.9) | $\mathrm{K}^2$ | what a reading will resolve | 7.2.1 |
| Last gain | $K_1$, eq. (7.12) | $\mathrm{K}^{-1}$ | the last step's law | 7.3.1 |
| First gain | $K_0$, eq. (7.15) | $\mathrm{K}^{-1}$ | one step more | 7.3.1 |
| Shrunk weight | $S/(1+S\sigma^2)$, eq. (7.18) | $\mathrm{K}^{-2}$ | what optimism does to a cost | 7.3.2 |
| Conditional violation | $v(y_1)$, eq. (7.20) | – | the risk after one reading | 7.3.3 |
| Overall violation | eq. (7.21) | – | the average of $v$ over the readings | 7.3.3 |
| Band as a cost on moving | eq. (7.22) | $\mathrm{K}^2$ | $s^2u^2$, where separation ends | 7.3.4 |
| Cautious move | eq. (7.25) | – | the band kept, learning left out | 7.3.4 |
| Exact move | eq. (7.26) | – | both terms of the recursion | 7.3.4 |

## 7.8 Exercises

**Exercise 7.1.** *Formulation.* A controller moves a drain valve to hold a tank's level reading at a set point, with the inflow unmeasured, and the tank must not overflow. Write the problem in the form of equation (7.1) and name its couplings. (The state holds the level and the inflow; $F$ is the mass balance with the valve's flow in it, $h$ reads the level, $\ell$ weighs level error and valve travel, and the limit is $\Pr(\text{level}_k \le \text{top for all } k \mid \mathbf{y}_{\text{now}}) \ge 1-\alpha$. Coupling 1: the valve is in the balance. Coupling 2: the level record comes from a plant the valve is moving, so inflow and valve are read together. Coupling 3: a move changes how well the flow through the valve is known if its coefficient is an entry. Coupling 4: the level's spread sets the back-off from the top.)

**Exercise 7.2.** *Derivation.* Derive $K_1$ from the last step. ($\rho u_1^2 + q(a\hat{x}_1 - \theta u_1)^2$ plus a constant in $P$, by the mean of a square. The derivative is $2\rho u_1 - 2q\theta(a\hat{x}_1 - \theta u_1) = 0$, so $u_1 = qa\theta\,\hat{x}_1/(\rho + q\theta^2)$, and $K_1 = 6.60386/105.56091 = 0.062560$ per K.)

**Exercise 7.3.** *Derivation.* Show from $\mathbb{E}[(\theta u)^2]$ why an unknown $\theta$ breaks separation, and give the last gain when $s = \hat{\theta}$. ($\mathbb{E}[(\theta u)^2] = u^2(\hat{\theta}^2 + s^2)$, so the expected square of $ax - \theta u + w$ holds $s^2u^2$, a term with the move in it; a band on $x$ gave only a constant. The derivative now gives $K_1(s) = qa\hat{\theta}/(\rho + q(\hat{\theta}^2 + s^2))$, with the band in the denominator. At $s = \hat{\theta}$, $s^2 = 20.14023$ and $K_1(s) = 0.03548$ per K.)

**Exercise 7.4.** *Derivation.* Show that leaving the noise free shrinks the cost weight, and give the shrunk weight at $\sigma = 0.50\,\mathrm{K}$. (Minimizing $q(m+w)^2 + w^2/\sigma^2$ over $w$ gives $w = -q\sigma^2 m/(1+q\sigma^2)$, and the minimum is $qm^2/(1+q\sigma^2)$. The weight is $q/(1+q\sigma^2) = 2.00000$ against $q = 4.00$, and the expectation of the same square keeps $q$ and adds the constant $q\sigma^2$. Carried through both steps, the first move is 0.05083, a ratio of 0.8065 to the law.)

**Exercise 7.5.** *Application.* Find the saving of the linear law over the fixed sequence at $s_0 = 0.5\,\mathrm{K}$. (For the fixed sequence, $\mathrm{sd}(x_2) = 0.12622\,\mathrm{K}$, the back-off is $0.20762\,\mathrm{K}$ and $\mathbf{g}^{\mathsf{T}}\mathbf{u}$

must reach 0.24295, so the cost is $\rho$ times the square of that over 22.86591, which the script gives as 0.06454 from the unrounded values. Minimizing over $u_0$, $c$ and $\kappa$ under the back-off, with $\mathrm{sd}(\hat{x}_1) = 0.18892\,\mathrm{K}$, gives 0.05204. The saving is 19.36 %.)

**Exercise 7.6.** *Application.* Under the linear law at $s_0 = 0.2\,\mathrm{K}$, find the conditional violation after a reading at +2 standard deviations, and say why it exceeds 0.05. (By equation (7.20), with $a - \theta\kappa = 0.09498$, the conditional mean of $x_2$ is $-0.07177 + 2 \times 0.09498 \times 0.09669$, the spread is $0.10402\,\mathrm{K}$, and the tail is 0.0701. It exceeds the level because the law leaves part of a high reading uncancelled: the slope 0.06081 is below $a/\theta = 0.08197$, and the average over the readings remains 0.05.)

# Chapter 8

# Stability Under Parameter Uncertainty

## 8.1 A faster loop on one number

The commissioning engineer wants a faster loop. The discharge controller runs at $K_p = 0.10$ per kelvin, and the commissioning sheet offers a way to raise it. Its table gives the critical gain of the rated coil at 20 kW, the lowest load it lists, as $K_c = 0.8223$; this is the proportional gain, with the ratio $K_i/K_p$ held, at which the sampled loop stops settling. The sheet applies its standard gain margin of 2 and writes $K_p = 0.4112$, which it rounds to 0.41 and calls safe.

The figure comes from the rated coil, and Chapter 5 has already overturned that coil. Chapter 6 left a band on the plant, and a steady window says nothing about the capacities that set how fast anything moves. The sheet has one number and one factor. It cannot say how likely the tuning is to be stable given the band, what the word stable delivers once it holds, or what the loop does if it does not hold. The question of this chapter is

*how likely this tuning is to be stable, given the band on the plant, and what stable buys.*

**Jargon: one-period map, eigenvalue, spectral radius, residue, critical gain, ultimate gain, gain margin, the controller's pair, limit cycle**

The *one-period map* is the matrix that carries the deviations of a sampled loop from one controller step to the next. An *eigenvalue* of the map is a number that a particular direction of the state is multiplied by at each step. The *spectral radius* is the largest modulus among the eigenvalues. The *residue* of a mode on a quantity is the amplitude that a given kick puts into that mode, read at that quantity. The *critical gain* is the gain at which the spectral radius reaches one, here along a stated line through the gain plane with the integrator in the loop. The *ultimate gain* usually means the same crossing under proportional action alone; on this plant, at the MAP state and 20 kW, that is 1.7236, not the 1.0129 of the line used here. A *gain margin* is the factor between the critical gain and the gain in use. *The controller's pair* (the dominant complex pole pair) is the complex eigenvalue pair that the gain moves and that carries the discharge air's response. A *limit cycle* is a steady oscillation of a nonlinear loop, whose size is set by the nonlinearity and not by the kick that began it.

**Key idea**

Stability of a sampled loop is an event on the state: the spectral radius of its one-period map below one. On this plant that event is one scalar root, so it is a threshold question on the critical gain. A steady window reads the loop's steady product very well and reads the capacities not at all, so a tight posterior can sit under a wide critical gain. The radius itself is not a measure of speed: a slow mode of the hall sets it, that mode barely moves above $K_p$ about 0.06, and a decay requirement has to be put on the mode the controller moves. Even then it is an envelope of that mode and not the settling of a load step, which the hall sets.

### 8.1.1 The spectral radius from a small example

The reader has met steady operating points so far, and a steady point says nothing about whether the plant gets there. The object that does is built here from the smallest loop, using the one-node plant of Section 7.1.5 with its values $a = 0.36788$ and $\theta = 4.48779\,\mathrm{K}$ per unit. These few lines are an illustration on that node, not the hall.

**One entry.** With a proportional law $u_k = Kx_k$ on the deviation $x_k$, equation (7.4) without its noise reads

$$x_{k+1} = (a - \theta K)\, x_k, \qquad x_k = (a - \theta K)^k\, x_0. \tag{8.1}$$

One number decides everything: the multiplier $a - \theta K$. Its modulus below one makes $x_k$ shrink to zero from any start, and its modulus above one makes $x_k$ grow from any start but zero. A gain that overshoots makes it negative, and $x_k$ then changes sign each step. Table 8.1 lists it in its upper part. Each row is a gain; the columns give the multiplier, its modulus and the first readings from $x_0 = 1\,\mathrm{K}$. The multiplier passes through zero at $K = a/\theta = 0.08197$ and reaches $-1$ at $K = (1 + a)/\theta = 0.30480$, the critical gain of this one-entry loop.

**Two entries.** Add an integrator, $u_k = Kx_k + I_k$ and $I_{k+1} = I_k + \beta K x_k$, with $\beta = 1.2$ the ratio $K_i/K_p$ times the step. The one number becomes a matrix on $\mathbf{z}_k = (x_k, I_k)^{\mathsf{T}}$,

$$\mathbf{z}_{k+1} = \mathbf{M}\,\mathbf{z}_k, \qquad \mathbf{M} = \begin{bmatrix} a - \theta K & -\theta \\ \beta K & 1 \end{bmatrix}, \qquad \mathbf{z}_k = \mathbf{M}^k \mathbf{z}_0. \tag{8.2}$$

A power of a matrix is read through its eigenvalues. If $\mathbf{M}\mathbf{v}_i = \lambda_i \mathbf{v}_i$ with independent eigenvectors, then write the start as $\mathbf{z}_0 = \sum_i c_i \mathbf{v}_i$. One application multiplies each term by its own $\lambda_i$, so

$$\mathbf{z}_k = \sum_i c_i\, \lambda_i^k\, \mathbf{v}_i. \tag{8.3}$$

Each eigenvalue is again a multiplier, one for each mode. A complex pair $|\lambda| \mathrm{e}^{\pm \mathrm{i}\phi}$ gives a real oscillation: the modulus $|\lambda|$ scales it at each step, and the angle $\phi$ turns it, with a period of $2\pi/\phi$ steps. All the terms die for every start exactly when every modulus is below one, which is the event on the largest of them. That largest modulus is the spectral radius,

$$\rho(\mathbf{M}) = \max_i |\lambda_i|, \qquad \mathbf{z}_k \to \mathbf{0} \text{ for every } \mathbf{z}_0 \iff \rho(\mathbf{M}) < 1. \tag{8.4}$$

For the $2 \times 2$ map the eigenvalues solve $\lambda^2 - (\operatorname{tr}\mathbf{M})\lambda + \det\mathbf{M} = 0$, and when they are complex their product is the squared modulus, $|\lambda|^2 = \det\mathbf{M} = a + \theta K(\beta - 1)$. At $K = 0.10$ the map is

$$\mathbf{M} = \begin{bmatrix} -0.0809 & -4.4878 \\ 0.12 & 1 \end{bmatrix}, \qquad \operatorname{tr}\mathbf{M} = 0.9191, \qquad \det\mathbf{M} = 0.4576,$$

so the pair is $0.4596 \pm 0.4964\mathrm{i}$ with modulus 0.6765. From $(1, 0)$ the deviation runs $-0.0809$, $-0.5320$, $-0.4519$, $-0.1719$, $+0.0488$, $+0.1235$, $+0.0912$, $+0.0273$: an oscillation that shrinks.

The radius governs the long run, not every step, since the size went up between the first step and the second. The lower part of Table 8.1 lists this map at six gains. The pair reaches modulus one at $K = (1 - a)/(\theta(\beta - 1)) = 0.70427$, the critical gain of this two-entry loop.

Table 8.1: The small example of equations (8.1) and (8.2), an illustration on Chapter 7's node with $\beta = 1.2$ stated. Upper part: one entry. Lower part: two entries.

| $K$ [1/K] | $a - \theta K$ [–] | Modulus [–] | $x_1$ [K] | $x_2$ [K] | $x_3$ [K] | $x_4$ [K] | $x_5$ [K] |
|---|---|---|---|---|---|---|---|
| 0.0000 | +0.3679 | 0.3679 | +0.3679 | +0.1353 | +0.0498 | +0.0183 | +0.0067 |
| 0.0500 | +0.1435 | 0.1435 | +0.1435 | +0.0206 | +0.0030 | +0.0004 | +0.0001 |
| 0.2000 | −0.5297 | 0.5297 | −0.5297 | +0.2806 | −0.1486 | +0.0787 | −0.0417 |
| 0.3000 | −0.9785 | 0.9785 | −0.9785 | +0.9574 | −0.9368 | +0.9166 | −0.8968 |
| 0.3500 | −1.2028 | 1.2028 | −1.2028 | +1.4468 | −1.7403 | +2.0933 | −2.5180 |

| $K$ [1/K] | Eigenvalues of $\mathbf{M}$ [–] | Moduli [–] | Radius $\rho$ [–] |
|---|---|---|---|
| 0.02 | +0.4888, +0.7893 | 0.4888, 0.7893 | 0.7893 |
| 0.05 | $+0.5717 \pm 0.2930$i | 0.6425, 0.6425 | 0.6425 |
| 0.10 | $+0.4596 \pm 0.4964$i | 0.6765, 0.6765 | 0.6765 |
| 0.30 | $+0.0108 \pm 0.7981$i | 0.7982, 0.7982 | 0.7982 |
| 0.70 | $-0.8868 \pm 0.4580$i | 0.9981, 0.9981 | 0.9981 |
| 0.75 | $-0.9990 \pm 0.2076$i | 1.0203, 1.0203 | 1.0203 |

**A slow entry that the loop does not feel.** The radius is a statement about all the entries together, and one more line of the example shows what that costs. Add a slow node $h_{k+1} = 0.95\,h_k + 0.05\,x_k$ that the loop's air warms and that does not feed back. The map is then block triangular, and the moduli at $K = 0.10$ are 0.9500, 0.6765 and 0.6765. The radius is 0.9500, and no gain can move it, since 0.95 is an entry of the map that $K$ does not touch. Yet $x_k$ has no part of that mode: its residue on $x$ in equation (8.3) is zero, because $h$ never reaches $x$. From $(1, 0, 0)$, the deviation $x_k$ is $+1.69 \times 10^{-5}$ by step 20 and $+9.67 \times 10^{-11}$ by step 60, while $h_k$ is $-1.87 \times 10^{-3}$ and $-2.37 \times 10^{-4}$. So the radius answers whether every entry settles. It need not say how fast the entry of interest settles, and the hall will show this in full.

### 8.1.2 The closed-loop hall as one matrix (step 0)

The plant is the closed-loop hall of Chapter 5 with the offset of Chapter 6, read through the one object above.

**The restriction.** This chapter works under the row "Fixed law, gains given" of Chapter 7's map, Table 7.6. In the problem (7.1) the policy line is filled in with the PI row, and no objective produced it; the first two couplings of Section 7.1.3 stay, and the third and fourth are given up. The chapter chooses nothing. A stated requirement below is a threshold, not a cost, and tuning as an objective is Chapter 14. The state is indexed by controller step $k$, 30 s apart; the time-stacked form waits for Chapter 9.

**The map.** Around an operating point at a pinned load $Q^*$ and set point $r$, with the valve off its stops, the three physical rows of the plant and the controller (5.1) give small deviations $\mathbf{x} = (\delta T,\ \delta T_c,\ \delta u)^{\mathsf{T}}$ and $\delta I$. Writing $u = Q^*/(W(T_c - T_w))$ and $T_c = r - b$ from equation (6.6),

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{A}\,\mathbf{x} + \mathbf{B}\,\delta c, \qquad \mathbf{A} = \begin{bmatrix} -G/C_h & G/C_h & 0 \\ G/C_c & -(G + Wu)/C_c & -W(T_c - T_w)/C_c \\ 0 & 0 & -1/\tau_a \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 1/\tau_a \end{bmatrix}. \tag{8.5}$$

The command is held for one step of $\Delta t = 30\,\mathrm{s}$, so one step is $\mathbf{x}_{k+1} = \mathbf{A}_d\mathbf{x}_k + \mathbf{B}_d\,\delta c_k$ with $\mathbf{A}_d = \mathrm{e}^{\mathbf{A}\Delta t}$ and $\mathbf{B}_d = \int_0^{\Delta t} \mathrm{e}^{\mathbf{A}s}\mathbf{B}\,\mathrm{d}s$. The controller adds $\delta c_k = K_p\,\delta T_{c,k} + \delta I_k$ and $\delta I_{k+1} = \delta I_k + K_i\Delta t\,\delta T_{c,k}$, and the four deviations then advance together:

$$\begin{bmatrix}\mathbf{x}_{k+1}\\ \delta I_{k+1}\end{bmatrix} = \mathbf{M}\begin{bmatrix}\mathbf{x}_k\\ \delta I_k\end{bmatrix}, \qquad \mathbf{M}(K_p;\mathbf{z}) = \begin{bmatrix}\mathbf{A}_d + K_p\,\mathbf{B}_d\,\mathbf{e}_2^{\mathsf{T}} & \mathbf{B}_d\\ K_i\Delta t\,\mathbf{e}_2^{\mathsf{T}} & 1\end{bmatrix}, \tag{8.6}$$

with $\mathbf{e}_2$ picking $\delta T_c$ out of $\mathbf{x}$. Stability of this tuning at this state is the event of equation (8.4) on $\mathbf{M}$. The sensor's offset does not appear as a deviation, since it is constant. It enters by $r \to r - b$ in the operating point: through $T_c - T_w$, in $u$ and in the loop's entry $W(T_c - T_w)$.

**The modeling choice that carries the chapter.** A steady window has no row on the entries of $\mathbf{A}$ that divide by a capacity or a time. The plant pins them at $C_c = 300\,\mathrm{kJ/K}$, $\tau_a = 60\,\mathrm{s}$ and $C_h = 3600\,\mathrm{kJ/K}$. Here they get independent log-normal priors, standard deviation 0.15 in the log, centered on the pinned values, which spans factors 0.7453 to 1.3418 at 95 %, about ±30 %. This buys a band that carries what the window cannot see, and it fails if the real coil's capacity is further off than that; nothing in the record corrects it, the priors' centers are the plant's own numbers, and Section 8.4.3 shows how far the width moves the answers. Figure 8.1 draws the loop. The boxes are the hall air, the coil with the supply air, the valve actuator and the sampled controller; solid arrows are the physical paths, and the two dashed notes say what the steady window reads and what it does not.

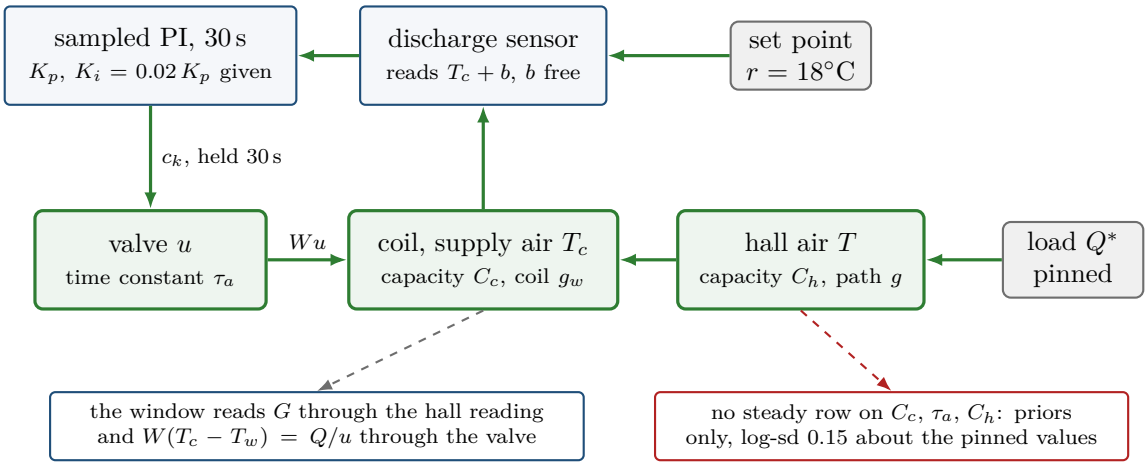



Figure 8.1: The closed-loop hall as the map (8.6) sees it. Green: the deviations of the one-period map, with the integrator's in the controller. Gray: pinned at the scenario node. The notes below say which entries of $\mathbf{A}$ the steady window reads (blue) and which it does not (red).

**The scenario node.** Every what-if but one is closed loop: the set point is pinned at $r = 18°\mathrm{C}$, the load at $Q^*$, and the valve's deviation stays free inside its stops, determined by the controller's row. The exception is a candidate test of step 7, a valve bump, which is open loop and pins the actuator. $K_i/K_p$ stays 0.02 throughout, so a gain is a point on one stated line through the $(K_p, K_i)$ plane.

### 8.1.3 The control engineer's questions (step 1)

Table 8.2 lists the seven questions. The first column is the question in the engineer's words, the second the general form, the third its mathematical form on this hall, the fourth the category, the fifth the computation and the last the subsection. The first and fifth read a variance of

a derived quantity; the second, third and fourth are one threshold read three ways; the sixth is a simulation; the last is the next measurement. Here $\lambda_{\text{pair}}$ is the controller's pair, which Section 8.2.4 defines on the map.

Table 8.2: The questions asked of the map. $\mathbf{y}$ is Chapter 6's three readings and three priors, with the three priors on the dynamic entries added.

| Control engineer's question | General form | Mathematical form | Category | Computation | Answered in |
|---|---|---|---|---|---|
| What does the band do to the critical gain? | moments of a derived quantity, (2.10) | $p(K_c(\mathbf{z}) \mid \mathbf{y})$, with $\rho(\mathbf{M}) = 1$ at $K_c$ | estimation of a derived quantity | posterior | 8.3.1 |
| How likely is this tuning to be stable? | threshold at a scenario node, (2.12), (6.1) | $\Pr(\rho(\mathbf{M}(K_p, \mathbf{z})) < 1 \mid \mathbf{y})$ | prediction under a threshold | posterior | 8.3.2 |
| What is the largest gain that holds 0.95? | the same, inverted, (6.17) | $K_{95}$: $\Pr(K_c > K_{95} \mid \mathbf{y}) = 0.95$ | prediction under a threshold | posterior | 8.3.3 |
| And if it must also decay? | the same, on the pair, (6.1) | $\Pr(\lvert\lambda_{\text{pair}}\rvert \le 0.8609 \mid \mathbf{y})$ | prediction under a threshold | posterior | 8.3.4 |
| Where does the band come from? | budget of a variance, (4.54) | $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ by entry | estimation of a derived quantity | posterior | 8.3.5 |
| What does instability look like with stops? | simulation at the truth | $\mathbf{x}_{k+1}$ of the nonlinear loop | prediction | – (simulation at the truth) | 8.3.6 |
| Which test would narrow it most? | next measurement, (2.16) | sd $K_c$ after each candidate | design of measurement | preposterior | 8.6 |

### 8.1.4 The evidence as rows (step 2)

The evidence is Chapter 6's, with three rows added. The readings are that chapter's means of the record's twenty minutes at about 30 kW, each an independent normal row given the state, and the priors on $g$, $g_w$ and $b$ are that chapter's, with the offset free. Those twenty minutes are steady, so no reading row contains a capacity. Write $\ell_c = \ln(C_c/300)$, $\ell_a = \ln(\tau_a/60)$ and $\ell_h = \ln(C_h/3600)$; the three new rows are their priors. With $*$ for an entry that can be nonzero, the Jacobian has the pattern

$$\mathbf{J} = \begin{bmatrix} * & * & 0 & * & 0 & 0 & 0 \\ * & 0 & * & * & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} \text{hall reading } y_T \\ \text{valve reading } y_u \\ \text{load meter } y_Q \\ \text{prior on } g \\ \text{prior on } g_w \\ \text{prior on } b \\ \text{prior on } \ell_c \\ \text{prior on } \ell_a \\ \text{prior on } \ell_h \end{matrix} \tag{8.7}$$

with columns $Q, g, g_w, b, \ell_c, \ell_a, \ell_h$. The fact to read here is the lower-right block: each of the last three columns has a line from its prior and from nothing else. So the fit cannot move these entries, and their spread after the record is their spread before it. The discharge reading is absent for the reason of Section 6.1.4: the loop consumes it.

## 8.2 The posterior of the state (step 3)

Every answer reads one posterior, Chapter 6's with three independent entries added. Once it is written down, the map can be computed at its center, and what the map does there fixes how the answers are read.

### 8.2.1 The rows and the covariance

The state is

$$\mathbf{z} = (Q,\ g,\ g_w,\ b,\ \ell_c,\ \ell_a,\ \ell_h)^{\mathsf{T}}. \tag{8.8}$$

The rows are the residual vector (6.9) with three more, $\ell_c/0.15$, $\ell_a/0.15$ and $\ell_h/0.15$, and the objective is $\mathcal{L} = \frac{1}{2}\mathbf{r}^{\mathsf{T}}\mathbf{r}$. The Jacobian is equation (6.12) with a diagonal block of 6.66667 added, by the pattern (8.7). The information matrix of these normalized rows is $\mathbf{J}^{\mathsf{T}}\mathbf{J}$; with unnormalized rows it is $\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$. The normal equations (6.10) separate into two blocks: the first four entries iterate exactly as in Table 6.2, and the last three are linear with their minimum at zero. The covariance, with the curvature term of equation (6.11) kept in its first block, is

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{\text{fit}} & \mathbf{0} \\ \mathbf{0} & 0.0225\,\mathbf{I}_3 \end{bmatrix}, \qquad \boldsymbol{\Sigma}_{\text{fit}} = \begin{bmatrix} 0.024957 & 0.000883 & 0.000762 & -0.00027 \\ 0.000883 & 0.001772 & -0.001827 & -0.007321 \\ 0.000762 & -0.001827 & 0.002348 & 0.00913 \\ -0.00027 & -0.007321 & 0.00913 & 0.036048 \end{bmatrix}, \tag{8.9}$$

at the MAP state $\hat{Q} = 30.0581\,\text{kW}$, $\hat{g} = 0.01833$, $\hat{g}_w = -0.16764$, $\hat{b} = 0.05708\,\text{K}$ and $\hat{\ell}_c = \hat{\ell}_a = \hat{\ell}_h = 0$. The spreads are 0.1580 kW, 0.04210, 0.04846 and 0.18986 K, with $\text{corr}(g_w, b) = 0.9924$, $\text{corr}(g, b) = -0.9160$ and $\text{corr}(g, g_w) = -0.8956$.

### 8.2.2 What the window reads

That ridge looks like bad news for a loop's gain, and it is not. The entry of the map that carries the valve's effect on the air is the product $W(T_c - T_w)$. At steady state the coil's balance makes it $Q/u$, the load over the valve, and both of those are read. So its log, $g_w + \ln(T_c - T_w)$, has the gradient $(0,\ 1,\ -1/(T_c - T_w))$ on $(g, g_w, b)$, and through $\boldsymbol{\Sigma}_{\text{fit}}$ its spread is 0.00596, against 0.04846 for $g_w$ alone and 0.04815 for $\ln(T_c - T_w)$ alone. The ridge of Chapter 6 runs along the direction in which that product stands still. The product is 53.350 kW per unit of valve at the MAP state, 53.566 in the stated truth and 64.000 for the rated coil.

### 8.2.3 The two routes

Two routes read the posterior, and every probability below is given by both, like with like. The *state route* linearizes the derived quantity at $\hat{\mathbf{z}}$ and uses the Laplace normal: a value, a gradient and $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$. The *draws route* takes Chapter 6's reweighted draws of the first four entries, a $t_5$ proposal with seed 8 and $N = 200\,000$, draws the dynamic entries independently from their priors with seed 9, and evaluates each event on the map itself. Its effective size is 119 928; its standard errors come from the weights and are given in parentheses. Each route comes in two bands. *Fit only* holds the dynamic entries pinned at their centers, and *fit + dynamic* carries them. Those are not two methods of one answer; they are different questions, since the first asks about a plant whose capacities are known.

### 8.2.4 The map at the MAP state

At the MAP state and 20 kW the operating point is $T = 21.2157°\mathrm{C}$, $T_c = 17.9429°\mathrm{C}$, $u = 0.3749$, with $G = 6.1110\,\mathrm{kW/K}$, $W = 13.5305\,\mathrm{kW/K}$ and $T_c - T_w = 3.9429\,\mathrm{K}$. Then

$$\mathbf{A} = \begin{bmatrix} -0.001697 & 0.001697 & 0 \\ 0.02037 & -0.037278 & -0.177832 \\ 0 & 0 & -0.016667 \end{bmatrix},$$
$$\mathbf{A}_d = \begin{bmatrix} 0.96105 & 0.0299 & -0.07994 \\ 0.35879 & 0.33435 & -2.42739 \\ 0 & 0 & 0.60653 \end{bmatrix}, \quad \mathbf{B}_d = \begin{bmatrix} -0.01523 \\ -0.80008 \\ 0.39347 \end{bmatrix}, \tag{8.10}$$

with $\mathbf{A}$ in $\mathrm{s}^{-1}$ and K/s per unit, and $\mathbf{A}_d$, $\mathbf{B}_d$ dimensionless and in K per unit; the third column of the second row is the loop's gain, negative because opening cools. At $K_p = 0.41$, where $K_i \Delta t = 0.246$,

$$\mathbf{M} = \begin{bmatrix} 0.96105 & 0.02365 & -0.07994 & -0.01523 \\ 0.35879 & 0.00631 & -2.42739 & -0.80008 \\ 0 & 0.16132 & 0.60653 & 0.39347 \\ 0 & 0.246 & 0 & 1 \end{bmatrix}. \tag{8.11}$$

This batched map agrees with the plant's own to a radius difference of 0 in the digits printed, also with $C_c$ and $\tau_a$ moved to 390 and 42.

**The modes.** Figure 8.2 shows what happens as the gain rises. The left panel is the complex plane with the unit circle: each marker is an eigenvalue of the map at one gain, with the pair in blue, the slow real mode in purple and the fast real mode in gray. The right panel gives the moduli against $K_p$: the radius, and the pair's modulus with the slow mode set aside. There are three real facts in it. The pair leaves the circle's interior as $K_p$ rises. The slow real mode stays put, near 0.95. And the fast real mode, between 0.3414 and 0.3884, is never in question.

The slow mode is the hall air. The discharge loop holds $T_c$, so the hall air relaxes toward it alone, through the path and its own capacity, with a time constant of $C_h/G = 589.1\,\mathrm{s}$ and a multiplier of $\mathrm{e}^{-\Delta t\, G/C_h} = 0.9503$. The eigenvector at $K_p = 0.41$ shows it, with the $T$ entry set to one: $(1,\ -0.0235,\ 0.1195,\ 0.1137)$, a hall-air movement with hardly any discharge air in it. The controller's pair is the valve and the coil, the two faster elements, at 60 s and 26.8 s at this load, with the integrator. To define it on any state: set aside the eigenvalue nearest $\mathrm{e}^{-\Delta t\, G/C_h}$, and $|\lambda_{\mathrm{pair}}|$ is the largest modulus among the rest. Its modulus reaches one at $K_p = 1.0129$ at the MAP state, and the angle at $K_p = 1.00$ gives a period of 5.34 steps.

**One scalar root.** On the line $K_i = 0.02\,K_p$ the radius crosses one once on $K_p \in [0.05, 20]$, by a grid of 600 points, at the MAP state and at four corners of $\pm 30\,\%$ on $C_c$ and $\tau_a$, near 1.015, 1.191, 0.823, 1.357 and 0.928. Across the draws at $K_p = 0.41$, 0.80 and 1.00 the event $\rho(\mathbf{M}) < 1$ and the event $K_c > K_p$ never disagree: the share is 0.000000 at each. So stability at this gain is one comparison, and the state route may go on the root $K_c(\mathbf{z})$ rather than on the radius, whose largest eigenvalue changes identity near 0.79. Bisection in $\ln K_p$ finds the root; 26 and 50 bisections give 1.01291582 and 1.01291584.

## 8.3 The answers (step 4)

Each answer is read from $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 8.2.1. The value reported is the state route's, with the draws route beside it where a probability is given, and the gap is said where it matters.

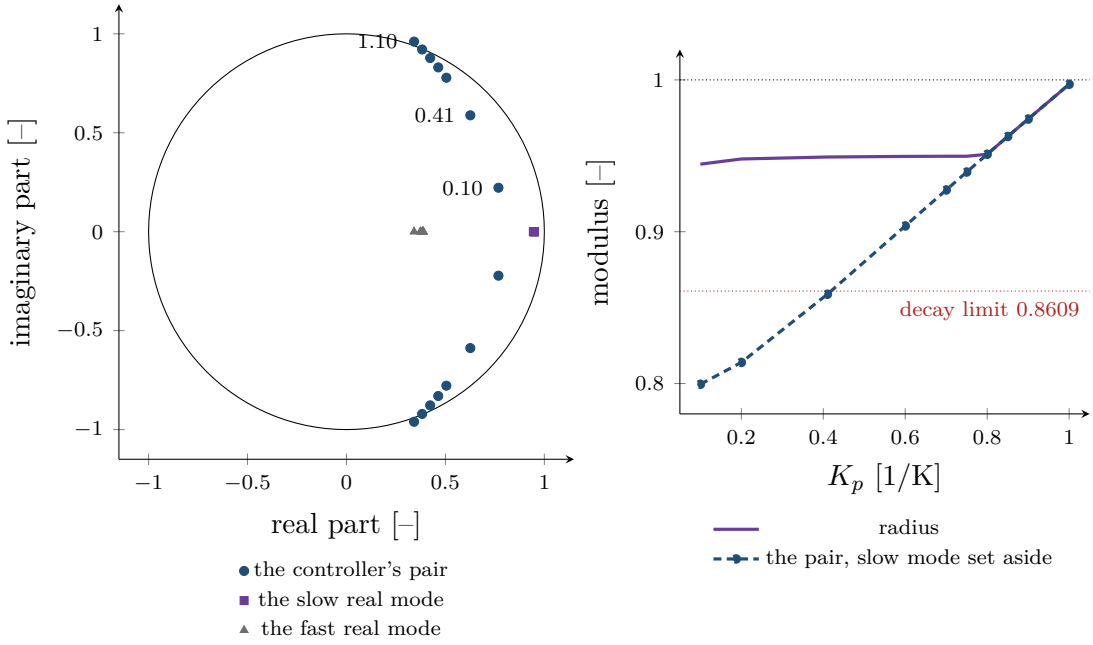



Figure 8.2: The eigenvalues of the map (8.6) at the MAP state, 20 kW. Left: at $K_p = 0.10$, 0.41, 0.70, 0.80, 0.90, 1.00 and 1.10, the pair labeled at three of them, with the unit circle. Right: the radius and the pair's modulus against $K_p$; the two lines meet near 0.79. Dotted: modulus one, and the decay limit of Section 8.3.4.

### 8.3.1 The critical gain under the band

**Control engineer's question.** What does the band do to the critical gain?

**General form.** Moments of a derived quantity under the posterior, equation (2.10), with the variance $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ of the delta method (4.53).

**Specialized form.** Equation (2.10) with $f = K_c(\mathbf{z})$, the root of $\rho(\mathbf{M}(K_p;\mathbf{z})) = 1$ on the map (8.6). Its gradient is taken numerically on the root at $\hat{\mathbf{z}}$, with a zero in the column of $Q$ because the node pins the load. The choice that enters is the line through the gain plane, $K_i/K_p = 0.02$, which makes $K_c$ a critical gain of that line and no other.

**Building blocks.** $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ of equation (8.9), and the gradient, at 20 kW,

$$\mathbf{g}_{K_c} = (0,\ 0.8717,\ -1.0129,\ 0.447,\ -0.625,\ 0.0118,\ 0.0165)^{\mathsf{T}}. \tag{8.12}$$

**The plant's numbers.** Table 8.3 gives $K_c$ by load. Each row is a load, and the columns give the rated coil's figure, the value at the MAP state, the fit-only spread by the state route, the stated truth and the truth's distance from the fit in that spread. With the dynamic entries carried, the spread at 20 kW is 0.0949 by the state route and 0.0983 by the draws (fit only: 0.0145 and 0.0152), and the draws' central 95 % runs from 0.8410 to 1.2303.

The weaker coil lowers the loop's gain, so the rated figure, 0.8223 against the fit's 1.0129, errs on the safe side here, for that reason. A coil that had turned out stronger than rated would have reversed it. Freeing the offset changes almost nothing here: the sensor-true fit of Chapter 5 gives 1.0124 with a spread of 0.0146. $K_c$ rises with load on every draw, with weight 1.0000 on $K_c(20) < K_c(30) < K_c(45)$. So 20 kW is the worst case only because it is the sheet's lowest load. It is the sheet's load, not the plant's: the rows at 10 and 15 kW show the fit's figure

Table 8.3: The critical gain by load, $K_i/K_p = 0.02$. The spread is fit only, state route. The truth's dynamic entries are the prior's center by construction, so the last column tests the fit, not the priors.

| $Q^*$ [kW] | Rated [1/K] | Fit [1/K] | sd, fit only [1/K] | Truth [1/K] | Truth to fit [sd] |
|---|---|---|---|---|---|
| 10 | 0.5390 | 0.6646 | 0.0178 | 0.6453 | −1.08 |
| 15 | 0.6752 | 0.8320 | 0.0145 | 0.8218 | −0.70 |
| 20 | 0.8223 | 1.0129 | 0.0145 | 1.0136 | +0.05 |
| 30 | 1.1478 | 1.4130 | 0.0286 | 1.4405 | +0.96 |
| 45 | 1.7120 | 2.1069 | 0.0667 | 2.1871 | +1.20 |

continuing down, 0.5116 at 5 kW, and the truth's $K_c$ at 2 kW is 0.398, under the sheet's gain, with the valve at 0.0373.

**The line through the gain plane.** $K_c$ belongs to one line. At the MAP state and 20 kW the ratio $K_i/K_p = 0.01$ gives 1.5588, the ratio 0.02 gives 1.0129, 0.03 gives 0.417 and 0.04 gives 0.2224; and $K_p$ raised alone with $K_i = 0.002$ held crosses at 1.7156. The sheet's "$K_p = 0.41$" therefore raises $K_i$ as well, to 0.0082, and its margin is gone at a 50 % change of the ratio: at 0.03 the critical gain is 0.417. Held $K_i$ is not free of cost either. At $K_p = 0.41$ with $K_i = 0.002$ the eigenvalues are 0.9388, 0.9089 and $0.3631 \pm 0.5205$i, so by the definition of Section 8.2.4, with the mode nearest the hall's set aside, the requirement of Section 8.3.4 fails at 0.9089. This chapter states that and does not optimize it.

### 8.3.2 The probability of stability

**Control engineer's question.** How likely is this tuning to be stable?

**General form.** The threshold form (2.12), in the form with a pinned input (6.1).

**Specialized form.** Equation (6.1), itself equation (2.12), with its input $r$ taken as $K_p$, $f(\mathbf{z}; K_p) = \rho(\mathbf{M}(K_p; \mathbf{z}))$ and $f^* = 1$, read as a strict event. By the zero disagreement of Section 8.2.4 it is the threshold $f = K_c$, $f^* = K_p$, and the state route is the normal tail (6.2) on the root,

$$\Pr\big(\rho(\mathbf{M}(K_p; \mathbf{z})) < 1 \mid \mathbf{y}\big) = \Pr\big(K_c(\mathbf{z}) > K_p \mid \mathbf{y}\big) \approx \Phi\left(\frac{K_c(\hat{\mathbf{z}}) - K_p}{\sqrt{\mathbf{g}_{K_c}^{\top} \boldsymbol{\Sigma}\, \mathbf{g}_{K_c}}}\right). \tag{8.13}$$

The draws route evaluates the same event on each draw's root.

**Building blocks.** $K_c(\hat{\mathbf{z}}) = 1.0129$, the gradient (8.12), $\boldsymbol{\Sigma}$ and $\Phi$. At $K_p = 0.80$, fit + dynamic, the argument is $(1.0129 - 0.80)/0.0949$ and the tail is 0.9876.

**The plant's numbers.** Table 8.4 gives the four routes at 20 kW. Each row is a gain, the middle columns the two routes in the two bands, with the draws' standard errors, and the last two columns what the rated coil's figure and the stated truth say. An entry >0.9999 is one whose draws' standard error is below the fourth decimal as well. Figure 8.3 draws these columns against $K_p$: solid for draws, dashed for the state route, purple for fit only and blue for fit + dynamic, with vertical lines at the sheet's 0.4112, the rated 0.8223 and, dotted, the truth's 1.0136. Points at >0.9999 are plotted at one.

Table 8.4: Pr(stable | **y**) at 20 kW, by the two routes in the two bands, with the draws' standard errors in parentheses. Draws: $t_5$ proposal, seed 8, $N = 200\,000$, effective size 119 928; dynamic entries seed 9.

| | Fit only | | Fit + dynamic | | | |
|---|---|---|---|---|---|---|
| $K_p$ [1/K] | State [–] | Draws [–] | State [–] | Draws [–] | Rated says | Truth |
| 0.41 | >0.9999 | >0.9999 | >0.9999 | >0.9999 | stable | stable |
| 0.70 | >0.9999 | >0.9999 | 0.9995 | 0.9997 | stable | stable |
| 0.80 | >0.9999 | >0.9999 | 0.9876 | 0.9918 (0.0002) | stable | stable |
| 0.85 | >0.9999 | >0.9999 | 0.9570 | 0.9685 (0.0005) | unstable | stable |
| 0.90 | >0.9999 | >0.9999 | 0.8829 | 0.9028 (0.0008) | unstable | stable |
| 0.95 | >0.9999 | >0.9999 | 0.7463 | 0.7683 (0.0012) | unstable | stable |
| 1.00 | 0.8130 | 0.8612 (0.0009) | 0.5541 | 0.5734 (0.0014) | unstable | stable |
| 1.05 | 0.0053 | 0.0163 (0.0004) | 0.3480 | 0.3651 (0.0014) | unstable | unstable |
| 1.10 | 0.0000 | 0.0000 | 0.1794 | 0.1992 (0.0012) | unstable | unstable |

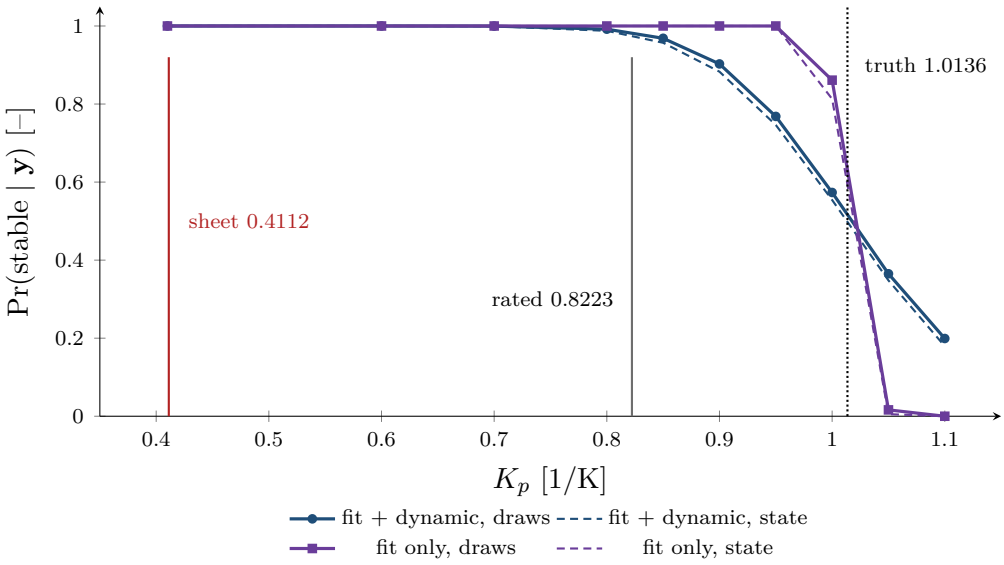



Figure 8.3: The probability of stability against the gain at 20 kW, from Table 8.4, with the values at 0.60 from the same computation. The dotted line is the stated truth's critical gain, whose dynamic entries are the prior's center by construction.

At the sheet's gain every route gives >0.9999. The two bands come apart as the gain rises: the fit-only curve holds at one until almost $K_c$ and then falls within 0.10 of gain, while the fit + dynamic curve begins to leave one near 0.70 and is about even at 1.00. By draws, $K_p = 0.85$ holds 0.9685 (0.0005), and the rated figure calls it unstable; that verdict is the rated coil's, and at this load it is conservative.

### 8.3.3 The largest gain that holds the level

**Control engineer's question.** What is the largest gain that holds 0.95?

**General form.** The chance constraint inverted for its input, equation (6.17), under the threshold form (2.12).

**Specialized form.** Equation (6.17), itself a reading of equation (2.12), with the input $K_p$ and the event of equation (8.13), so that $K_{95}$ is the 5 % point of $K_c$ under the posterior, $\Pr(K_c > K_{95} \mid \mathbf{y}) = 0.95$. The state route is the normal form (6.3), $K_{95} = K_c(\hat{\mathbf{z}}) - 1.6449\sqrt{\mathbf{g}_{K_c}^{\top}\,\boldsymbol{\Sigma}\,\mathbf{g}_{K_c}}$, and the draws route reads the weighted 5 % quantile.

**Building blocks.** $K_c(\hat{\mathbf{z}})$, the spread of Section 8.3.1 and, for the draws, the weights.

**The plant's numbers.** At 20 kW, fit only: 0.9890 by the state route and 0.9919 by the draws (se 0.0001). Fit + dynamic: 0.8568 by the state route and 0.8693 by the draws (se 0.0005; another seed, 18 and 19, gives 0.8699). By the draws, fit + dynamic, $K_{95}$ is 1.2142 at 30 kW and 1.7864 at 45 kW.

The fit alone would answer 0.99 with a spread of 0.015. A coil with $C_c$ only 15 % over the pinned value is unstable at that gain, with radius 1.0137, and at 30 % over, 1.0274; moved ±30 %, $\tau_a$ stays inside, at 0.9950 and 0.9926. The fit-only figure is not a tighter method beside the full one. It is the same question with entries pinned, so its larger $K_{95}$ is not a gain of any kind: its guarantee says nothing of a coil whose capacity differs.

At the sheet's end, no draw of 199 999 has $K_c$ at or below 0.4112, and the state route places it 6.34 standard deviations out. That distance is well past where the normal has been tested, and the draws can show a zero only down to about one part in their effective size, so the statement is "far below any level asked about", not a computed tail. The rated 0.8223 sits at $\Pr(K_c > 0.8223 \mid \mathbf{y}) = 0.9848$ (0.0003) by the draws. For stability alone the sheet is safe, with room, on these priors, at this load, and with no sensor lag; Section 8.6.5 shows how far each of those conditions carries.

### 8.3.4 The decay requirement

**Control engineer's question.** And if it must also decay?

**General form.** The threshold form (2.12) at a pinned input again, equation (6.1), and its inversion (6.17).

**Specialized form.** Equation (2.12) needs a quantity that measures speed. The engineer raised the gain to get a faster loop, so the natural step is to read speed from the radius. That does not work. From $K_p \approx 0.10$ to 0.79 the radius is the hall air's mode, 0.9445 to 0.9497, much as in the slow node of Section 8.1.1, though here the mode does feel the gain at the low end. At $K_p = 0.0442$, where the radius is smallest, the hall's mode and the integrator's have merged into a complex pair, $0.9047 \pm 0.0075\mathrm{i}$; above that the mode separates and runs from 0.9369 at $K_p = 0.06$ to 0.9499 at 1.10, so it barely moves above $K_p$ about 0.06. After a unit kick on the

discharge air at $K_p = 0.41$, the residue of that mode on $T_c$ is 0.0008 K, against 0.5494 K for the pair; on $T$ the two are comparable, 0.0338 and 0.0368 K. So the mode is a real movement of the hall air, and after a kick on $T_c$ the discharge error hardly contains it. Another kick is different, as below. Judged by the radius the best gain would be 0.0442, with a radius of 0.9047, and a loop at $K_p = 0.10$ with radius 0.9445 would count as faster than one at 0.41 with radius 0.9492. A radius set by the hall air is a fact about the hall air.

The requirement is therefore put on the controller's pair of Section 8.2.4. A stated requirement, an envelope of the pair down to 5 % in 10 minutes or 20 steps, gives $|\lambda_{\text{pair}}| \le 0.05^{1/20} = 0.8609$. This is a choice, and it is varied below. Equation (6.1) then has $f = |\lambda_{\text{pair}}(K_p; \mathbf{z})|$ and $f^* = 0.8609$. The pair's modulus rises with $K_p$ over the working range, though not everywhere: at the fit it is 0.7986, 0.7982 and 0.7995 at $K_p = 0.06$, 0.08 and 0.10, and on 13.3 % of the weight it falls somewhere on 0.10 to 0.5. So the event is treated as one root, and the bisection is checked against the map below: $K_p \le K_d(\mathbf{z})$, with $K_d$ the largest gain at which the requirement holds, and $K_d = 0$ for the states in which 0.10 already fails. Gains below 0.10 are not examined. The state route is the normal tail (6.2) on $K_d$, and the draws route takes the same event directly on the map's pair.

The event is an envelope of the pair and it is not the speed the gain was raised for. Its modulus rises over the working range, so raising $K_p$ cannot improve this envelope there; the event can only be lost as the gain goes up. What the engineer would see is in Table 8.5, the nonlinear loop at the fit, noiseless, at 20 kW. Each row is a gain; the columns give the pair's modulus and the steps until the discharge error stays within 5 % of its peak after a 10 kW load step and after a set-point step of +0.25 K, with the step itself as the peak.

Table 8.5: The settling of the discharge error at the fit, 20 kW, against the pair's modulus: steps to stay within 5 % of the peak.

| $K_p$ [1/K] | Pair's modulus [–] | 10 kW load step [steps] | Set-point step [steps] |
|---|---|---|---|
| 0.10 | 0.7995 | 59 | 17 |
| 0.20 | 0.8139 | 54 | 19 |
| 0.41 | 0.8588 | 48 | 22 |
| 0.60 | 0.9038 | 45 | 37 |

A load step moves the hall air by itself, so it fills the hall's mode. At $K_p = 0.41$ its error at step 20 is still 23 % of the peak, and the tail falls by about 0.945 per step, which is the hall's mode. Raising the gain buys a little on load steps, where the hall sets the time, and costs on set-point steps, where the pair sets it.

**Building blocks.** $K_d(\hat{\mathbf{z}}) = 0.4188$; its gradient in the log, 1.392, −1, 0.572, −1.713, −1.247 and 0.065 on $(g, g_w, b, \ell_c, \ell_a, \ell_h)$; $\mathbf{\Sigma}$; and hence spreads of 0.0099 fit only and 0.1335 fit + dynamic. Here $\tau_a$ counts, unlike in $K_c$.

**The plant's numbers.** Table 8.6 gives the probability against the gain. Each row is a gain; the columns give the two routes in the two bands, with the pair's modulus at the fit and the truth. The event on $K_d$ and the event on the map's pair agree in all four decimals at every gain, by the draws.

At the sheet's 0.41 the requirement holds with probability 0.5263 by the state route and 0.5309 (se 0.0014) by the draws. In the fit-only band it is 0.8134 and 0.8609. The stated truth holds it, with a pair of 0.8581, and that row is not a test of $C_c$ or $\tau_a$, since those sit at the prior's center. The two guarantees in this chapter are different events and are not to be ranked against each other: $K_{95} = 0.8693$ is a level on stability and 0.1858 is a level on a chosen speed. On 1.48 % of the weight the requirement fails even at 0.10. The sheet's tuning is almost certainly

Table 8.6: $\Pr(|\lambda_{\text{pair}}| \le 0.8609 \mid \mathbf{y})$ at 20 kW, and the pair's modulus. Below, the largest gain holding the requirement at 0.95. The figure 0.8609 in the fit-only column at 0.41 equals the requirement's number by coincidence.

| | Fit only | | Fit + dynamic | | Pair's modulus | |
|---|---|---|---|---|---|---|
| $K_p$ [1/K] | State [–] | Draws [–] | State [–] | Draws [–] | Fit [–] | Truth [–] |
| 0.10 | 1.0000 | 1.0000 | 0.9915 | 0.9852 (0.0003) | 0.7995 | 0.7980 |
| 0.20 | 1.0000 | 1.0000 | 0.9494 | 0.9395 (0.0007) | 0.8139 | 0.8128 |
| 0.30 | 1.0000 | 1.0000 | 0.8133 | 0.8056 (0.0011) | 0.8340 | 0.8331 |
| 0.35 | 1.0000 | 1.0000 | 0.6969 | 0.6938 (0.0013) | 0.8451 | 0.8443 |
| 0.41 | 0.8134 | 0.8609 | 0.5263 | 0.5309 (0.0014) | 0.8588 | 0.8581 |
| 0.45 | 0.0008 | 0.0044 | 0.4077 | 0.4136 (0.0014) | 0.8682 | 0.8675 |
| 0.50 | 0.0000 | 0.0000 | 0.2716 | 0.2735 (0.0013) | 0.8800 | 0.8794 |
| $K_d$ at 0.95 | 0.4025 | 0.4045 | 0.1992 | 0.1858 | $K_d$: fit 0.4188, truth 0.4219 | |
| At $K_p = 0.41$, draws, by load: 30 kW 0.9520; 20 kW 0.5309; 15 kW 0.1606; 10 kW 0.0104; 5 kW 0.0000 | | | | | | |

stable and, on these priors, at 20 kW and with no sensor lag, about even odds of this envelope. The last row of the table shows that 20 kW is the sheet's load and not the plant's: the same event at 0.41 holds with 0.9520 at 30 kW and 0.0104 at 10 kW. At 5 kW, $K_{95}$ is 0.4198 and the probability of stability at 0.41 is 0.9651.

The requirement moves this, and it is a choice. At 0.41, fit + dynamic, draws: 7.5 minutes ($|\lambda_{\text{pair}}| \le 0.8190$) holds with 0.1174, 10 minutes with 0.5309, 15 minutes (0.9050) with 0.9502 and 20 minutes (0.9278) with 0.9954. Five minutes (0.7411) fails at the fit and at the truth from $K_p = 0.10$ up.

### 8.3.5 Where the band comes from

**Control engineer's question.** Where does the band come from?

**General form.** The budget of a variance, equation (4.54), under the moments (2.10).

**Specialized form.** Equation (2.10) with $f = K_c$, its variance split by equation (4.54) with the gradient (8.12) and the blocks of $\mathbf{\Sigma}$ in equation (8.9). Since those blocks are independent, the budget separates into a term from the fit, with its cross terms, and one term per dynamic entry. The whole of it is first order, at the fit.

**Building blocks.** The gradient and $\mathbf{\Sigma}$. The fit block's terms split further by the map: the logs of $K_c$ have gradient $(0.8606,\ -1,\ 0.4413)$ on $(g, g_w, b)$, and this is nearly a sum of two lines the record reads. It is $-1$ times the coil's line $(0,\ 1,\ -0.2536)$ of Section 8.2.2, $-0.1750$ times the hall reading's line $(-4.9187,\ 0,\ -1)$ from Section 6.2.1, and a remainder of 0.0127 on $b$ alone.

**The plant's numbers.** Table 8.7 gives the budget at three loads, and Figure 8.4 draws the first of them. In the upper panel of the figure each horizontal bar is one term of the variance at 20 kW in $(1/\text{K})^2$, on a single scale, with the cross term negative. The two lower panels give $K_c$ against $\tau_a$ alone at 20 kW and 45 kW, the horizontal axis in factors of the pinned value.

The fit is 2.3 % of the variance at 20 kW by the state route (2.4 % by the draws). The unseen share is 97.7 %. That figure rests on the stated priors, whose center is the plant's own value, and it is this example's; at a prior of 0.05 in the log it is 82.2 % (Table 8.10). It is almost all $C_c$, whose gradient in the log is $-0.617$: a coil with more capacity is slower, and so has

Table 8.7: The budget of $\mathrm{Var}(K_c)$ by the state route, in $(1/\mathrm{K})^2$; gradients at the fit, prior log-sd 0.15 on the dynamic entries. The shares rest on that prior, whose center is the plant's own value.

| Term $[(1/\mathrm{K})^2]$ | 20 kW | 30 kW | 45 kW |
|---|---|---|---|
| $g$ | 0.001346 | 0.001743 | 0.002432 |
| $g_w$ | 0.002409 | 0.004688 | 0.010424 |
| $b$ | 0.007203 | 0.016757 | 0.044070 |
| Cross terms | −0.010747 | −0.022373 | −0.052490 |
| Fit total | 0.000211 | 0.000815 | 0.004436 |
| $C_c$ | 0.008789 | 0.017070 | 0.040133 |
| $\tau_a$ | 0.000003 | 0.001025 | 0.007993 |
| $C_h$ | 0.000006 | 0.000006 | 0.000007 |
| Total | 0.009009 | 0.018916 | 0.052570 |
| Unseen share, state route [%] | 97.7 | 95.7 | 91.6 |
| Fit's share, state route; draws [%] | 2.3; 2.4 | 4.3; 4.1 | 8.4; 7.9 |

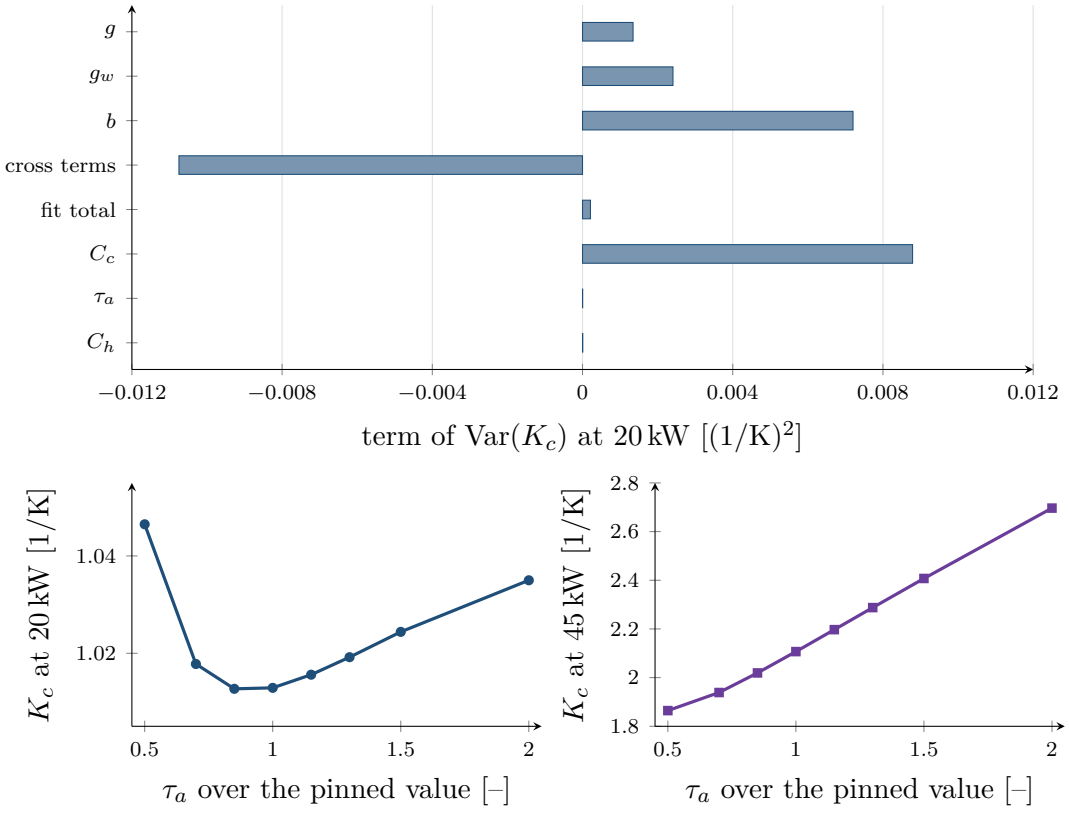



Figure 8.4: Upper: the terms of the budget at 20 kW, from Table 8.7. The three of the fit and the cross term nearly cancel to the bar marked fit total; $C_c$ does not cancel with anything. Lower: $K_c$ at the fit with $\tau_a$ alone moved, showing the minimum at 20 kW that a first-order budget reads as zero, and the monotone rise at 45 kW.

less room. The fit's own part is small through cancellation. The offset's term is the largest in the fit and the coil's follows, but the cross terms, $+0.003226$ for $(g, g_w)$, $-0.005705$ for $(g, b)$ and $-0.008268$ for $(g_w, b)$, remove 98.1 % of the sum of the three. This is Chapter 6's ridge read through the split above: on $\mathbf{\Sigma}_{\text{fit}}$ the coil's line has a spread of 0.00596, the hall line of 0.08307 K with its multiplier of $-0.1750$, and the remainder times the spread of $b$ is 0.00242; with their correlations these make the fit-only 0.01434 in $\ln K_c$. It also depends on the prior, as in Chapter 6: at a prior of 0.80 K on $b$ the fit-only spread is 0.0319, and the values of that sensitivity at 0.05, 0.10, 0.20 and 0.40 K are 0.0146, 0.0145, 0.0145 and 0.0175.

**A term that the first order misreads.** At 20 kW the budget gives $\tau_a$ almost nothing, 0.000003, which is a surprise for an element of the loop that has the longest time constant after the hall air. The reason is in the left lower panel: $K_c$ has a minimum between 0.7 and 1 times the pinned value, where the gradient vanishes, and it rises on both sides. It is 1.0465, 1.0178, 1.0127, 1.0129, 1.0192 and 1.0350 at factors 0.5, 0.7, 0.85, 1, 1.3 and 2, with the smallest value on a fine grid at factor 0.91. So the state route's zero is local, and the draws carry the rest. The first order also misses a joint effect. With $C_c$ at 1.3 times the pinned value, $\tau_a$ at factors 0.7 and 1.3 gives 0.9247 and 0.8197; with $C_c$ at 0.7 times, 1.1814 and 1.3512. So $\tau_a$ matters at 20 kW once $C_c$ is off center. In the exact decomposition of the variance under these two priors the interaction is 0.000266, larger than the whole fit's 0.000211, and $\tau_a$'s main effect is 0.000009. At 45 kW the curve is monotone, with $\mathrm{d}\ln K_c/\mathrm{d}\ln\tau_a = +0.2829$: a slower actuator raises $K_c$ there, against the intuition that a lag only hurts. On this line through the gain plane, at these loads, this is what the computation shows, and it is not a general property of actuators.

### 8.3.6 What instability looks like with stops

**Control engineer's question.** What does instability look like with stops?

**General form.** This question has no integral over the state. It is the simulation of the nonlinear loop, at the stated truth, with the plant's own integration, the PI row (5.1) and its conditional integration.

**The modeling choice.** A simulation at the truth gives one plant's behavior, so no probability attaches to anything in this subsection. The map's radius applies to the unsaturated loop only; here the valve's stops are in the loop. The load is fixed and there is no noise. At 20 kW the truth's $K_c$ is 1.0136 and its valve sits at 0.3734.

**The plant's numbers.** Table 8.8 gives the settled behavior: the last 120 of 1200 steps. Each row is a gain as a share of $K_c$ and a size of kick; the columns give the peak-to-peak of the discharge air, the period, the range of the valve and of the command, the mean error of the reading, the peak-to-peak of the hall air, the share of steps with the command at a stop, and the valve's travel per hour. Figure 8.5 draws the record: the upper left panel has $T_c$ against step over the last 30 steps, at 0.98 and $1.02\,K_c$; the upper right the valve and the command at $1.02\,K_c$ with the stops dotted; the lower left the peak-to-peak of $T_c$ in blocks of 60 steps, on a logarithmic axis, for three runs; and the lower right the hall air on its own scale.

There is no divergence, and the valve itself never reaches a stop at 20 kW, which a picture of instability as a valve slamming between its stops would not suggest. The command clips, and the actuator's 60 s filters the clipped signal, so the valve ranges over 0.22 to 0.54. At $1.02\,K_c$ the discharge air swings 0.90 K peak to peak with a period of 5.2 steps, or 2.6 minutes; the command reaches a stop in 1 to 2 % of steps; the hall air moves 0.046 K; and the valve travels about 12 per hour.

Table 8.8: The nonlinear loop at the stated truth, 20 kW, last 120 of 1200 steps, noiseless. Peak-to-peak is abbreviated p-p.

| $K_p/K_c$ [–] | Kick [K] | $T_c$ p-p [K] | Period [steps] | $u$ [–] | $c$ [–] | Mean error [K] | $T$ p-p [K] | At a stop [–] | Travel [1/h] |
|---|---|---|---|---|---|---|---|---|---|
| 0.98 | 2.00 | 0.0045 | 5.35 | 0.3726–0.3741 | 0.372–0.375 | −0.0000 | 0.0002 | 0.00 | 0.05 |
| 1.02 | 0.01 | 0.9014 | 5.23 | 0.2237–0.5382 | −0.001–0.773 | −0.0024 | 0.0460 | 0.01 | 12.15 |
| 1.02 | 2.00 | 0.9102 | 5.20 | 0.2222–0.5397 | −0.004–0.777 | −0.0056 | 0.0463 | 0.02 | 12.55 |
| 1.10 | 2.00 | 0.9816 | 5.02 | 0.1998–0.5601 | −0.039–0.866 | −0.0186 | 0.0552 | 0.05 | 12.38 |
| 1.50 | 0.01 | 0.4984 | 4.00 | 0.2655–0.4898 | −0.043–0.716 | −0.0827 | 0.0157 | 0.25 | 13.31 |
| 1.50 | 2.00 | 0.8137 | 4.30 | 0.2145–0.5803 | −0.260–0.833 | −0.0649 | 0.0322 | 0.15 | 15.55 |
| 2.00 | 2.00 | 0.7763 | 4.00 | 0.2067–0.5905 | −0.431–1.125 | −0.1396 | 0.0323 | 0.47 | 20.20 |

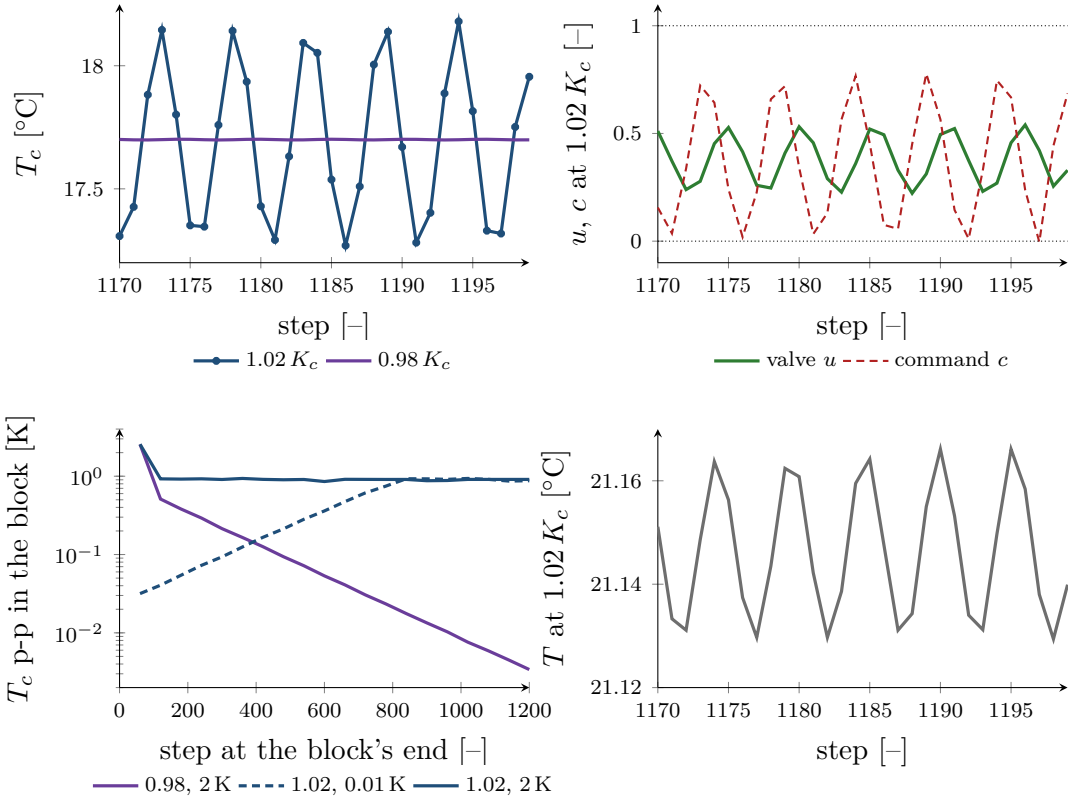



Figure 8.5: The nonlinear loop at the stated truth, 20 kW, one side of the critical gain and the other. Upper left: the discharge air in the last 30 of 1200 steps after a 2 K kick. Upper right: valve and command at 1.02 $K_c$, stops dotted: the command touches a stop and the valve does not. Lower left: peak-to-peak of $T_c$ per block of 60 steps. Lower right: the hall air, on its own scale.

The lower left panel shows the two sides. Below $K_c$ a 2 K kick dies, down to 0.0045 K in the last 120 steps at $0.98\,K_c$. Just above it any kick reaches about 0.9 K: the 0.01 K kick grows for hundreds of steps and arrives at the same size. The onset is abrupt in amplitude in this noiseless run, since the unsaturated map is linear and sets no size. With sensor noise it is not: with 0.10 K of white noise on the reading, on one seed, $T_c$ swings 1.22 K peak to peak at $0.98\,K_c$ and 1.16 K at $1.02\,K_c$, so the noise-driven swing hides the crossing. The size comes from where the nonlinearity begins, at the first clip, with the command's excursion $K_p$ times the error reaching the valve's distance to a stop, $\min(u, 1-u)/K_p = 0.3611$ K at $1.02\,K_c$, the order of half the peak-to-peak.

Further out, this is a nonlinear loop and not one number. At $1.5\,K_c$ two cycles coexist, 0.50 and 0.81 K, selected by the kick. At $2\,K_c$ the mean reading error reaches $-0.14$ K, so the loop no longer holds its set point on average. At 45 kW and $1.10\,K_c$ the cycle is 0.196 K, with the command clipping at 1 (its largest value 1.019) and the valve between 0.7886 and 0.8866. Stated plainly, on this example instability costs valve travel, not the hall's limit. The hall air barely sees it, and nothing here gives that a probability or claims it elsewhere.

## 8.4 The checks (step 5)

Six tests stand against the answers: the linear map against the nonlinear loop, the normal approximation near the boundary, the prior the record does not correct, the numerical settings, what the gain is seen to buy, and the stated truth. Table 8.9 collects the first and fourth; the others have their own subsections.

### 8.4.1 The map against the nonlinear loop and the numerics

Table 8.9 sets each check against its value. Each row is one check; the columns name the check, what it compares and the result. In its first rows a noiseless kick of 0.001 K runs in the plant's own integration at the truth, and its measured growth per step is compared with the pair's modulus on the map.

Table 8.9: The numerical checks of the map and of the state.

| Check | Compared | Result |
|---|---|---|
| Map's pair against nonlinear growth, at $0.90\,K_c$ | modulus against growth per step | 0.97681; 0.97710 |
| The same at $0.98\,K_c$ | the same | 0.99539; 0.99536 |
| The same at $1.02\,K_c$ | the same | 1.00459; 1.00439 |
| The same at $1.10\,K_c$ | the same | 1.02279; 1.02280 |
| Batched map against the plant's | radius difference, also at $C_c = 390$, $\tau_a = 42$ | 0 in the digits printed |
| Bisection | 26 against 50 steps, $K_c$ at the fit | 1.01291582; 1.01291584 |
| Integration step of the nonlinear loop | $T_c$ p-p at $1.10\,K_c$, 1 against 0.25 s | 0.99216 K both |
| $\sigma_p \to 0$ on the seven-entry full state of Section 6.4.1 | $K_c$ (fit-only sd) at $\sigma_p = 0.1$, 0.01, 0.001 | 1.0077 (0.0487); 1.0129 (0.0153); 1.0129 (0.0145) |

The map predicts the nonlinear loop's growth to about the fourth decimal on both sides of the critical gain. A radius computed on deviations therefore says what it should about the unsaturated loop. The reduced state reaches the full state's values by $\sigma_p = 0.001$.

### 8.4.2 The normal approximation near the boundary

A normal approximation is doing one specific job here: it converts a spread on $K_c$ into a tail, and the tail asked about is near the boundary, where the error sits. Beside the draws, the state

route's $K_{95}$ holds 0.9630 in the fit + dynamic band and 0.9675 fit only. So at the 5 % point it is acceptable and conservative. The fit-only probabilities differ by about 0.05 near 1.00 (0.8130 against 0.8612), which is Chapter 6's non-normal posterior, with a draws mean of 1.0163 against the fit's 1.0129; and the log-normal of $C_c$ skews the rest. That is why the draws' fit + dynamic mean is 1.0217.

The approximation is worse on the radius itself. Linearizing $\rho$ near the mode switch asks one slope to speak for two eigenvalues. At $K_p = 0.80$, with radius 0.9510 and a spread of 0.0243 by the gradient, that state route reads 0.9781 against the draws' 0.9918; at 0.75, where the slow mode sets the radius, its spread 0.0077 reads 1.0000 against 0.9984. This is why the state route goes on $K_c$.

### 8.4.3 The prior on the dynamic entries

The record cannot correct these priors. Table 8.10 varies them together. Each row is a prior, the columns the spread of $K_c$ and $K_{95}$ by the draws, and the unseen share by the state route.

Table 8.10: The prior on the dynamic entries at 20 kW, fit + dynamic. The stated choice is bold. The uniform row is ±30 %.

| Prior, log-sd [–] | sd $K_c$, draws [1/K] | $K_{95}$, draws [1/K] | Unseen share, state [%] |
|---|---|---|---|
| 0.05 | 0.0350 | 0.9605 | 82.2 |
| 0.10 | 0.0655 | 0.9152 | 94.9 |
| **0.15** | 0.0983 | 0.8693 | 97.7 |
| 0.20 | 0.1333 | 0.8236 | 98.7 |
| 0.30 | 0.2123 | 0.7323 | 99.4 |
| Uniform | 0.1169 | 0.8693 | – |

Even at the narrowest prior the unseen share is 82.2 %. $K_{95}$ falls steadily as the prior widens, and it is a figure the commissioning log has to state with its prior. The uniform's $K_{95}$ equals the normal's 0.8693 with a larger spread: this is a coincidence of the tails, and it should not be read as a robustness result.

### 8.4.4 What the gain buys

Stable and fast are not the whole of it. Table 8.11 runs the truth at six gains, with each row a gain. The peak error is from a noiseless 10 kW load step; the spread of $T_c$ and the valve's travel are from one noisy record of the record's noises, seed 81, 480 steps, one draw and not an average.

Table 8.11: The truth at 20 kW: what a higher gain gains and costs.

| $K_p$ [1/K] | Peak error, load step [K] | sd $T_c$, one record [K] | Travel, one record [1/h] | Pair's modulus [–] |
|---|---|---|---|---|
| 0.10 | 0.1410 | 0.0421 | 0.414 | 0.7980 |
| 0.20 | 0.0941 | 0.0621 | 0.906 | 0.8128 |
| 0.41 | 0.0614 | 0.0999 | 2.195 | 0.8581 |
| 0.60 | 0.0503 | 0.1348 | 3.798 | 0.9033 |
| 0.80 | 0.0394 | 0.2233 | 7.733 | 0.9507 |
| 1.00 | 0.0367 | 0.2517 | 9.966 | 0.9969 |

From 0.10 to 0.41 the load-step peak falls by more than half, while the spread of the air and the valve's travel rise, because the gain feeds the sensor's noise to the valve. With this sensor, faster costs noise. This is reported and not weighed; weighing it is Chapter 14.

### 8.4.5 The stated truth

At 20 kW the truth's $K_c$ is 1.0136, inside the band at +0.05 of the fit-only spread. This row flatters. The truth's dynamic entries are the prior's center by construction, so it tests the fit's product and path, and it tests nothing about $C_c$, $\tau_a$ or $C_h$. The dotted marker in Figure 8.3 is to be read the same way. A record with $C_c$ away from 300 would put the truth elsewhere in the fit + dynamic band, and this example holds no such record.

## 8.5 The report (step 6)

On the stated priors, with no sensor lag, the sheet's $K_p = 0.41$ is stable with room at 20 kW: no draw of 199 999 has a critical gain below it. It holds a 10-minute envelope of the pair with about even odds, 0.5263 by the state route and 0.5309 (se 0.0014) by the draws. Neither figure travels far. A discharge sensor lag of 5 s, which the band does not hold, takes the decay to 0.0453 and one of 20 s takes stability to 0.6052; at 10 kW the decay is 0.0104, and at 5 kW stability is 0.9651. The envelope is not the settling either: raising the gain buys a little on load steps, set by the hall, and costs on set-point steps. And on the stated priors, whose center holds the truth by construction, 97.7 % of the band on the critical gain comes from what the window cannot see; this percentage is the example's. The fit alone would allow 0.9919 where the full band allows 0.8693, both by the draws (0.9890 and 0.8568 by the state route); those are different questions, not a tighter and a looser method, and the fit-only gain is unstable for a coil of 15 % more capacity.

### 8.5.1 The answers and their bands

Table 8.12 gives the answers in the order of Table 8.2. Each row is a question; the second column gives the state-route answer with the draws beside it, the third the band, and the last the stated truth.

### 8.5.2 The entry in the commissioning log

The entry records the tuning and the load: $K_p = 0.41$, $K_i = 0.0082$, at 20 kW, with the set point at 18°C. It records Pr(stable) > 0.9999 by both routes and Pr(decay to 5 % in 10 minutes) = 0.5263 by the state route and 0.5309 by the draws. It declares the priors on $C_c$ and $\tau_a$ as assumptions, log-sd 0.15 about 300 kJ/K and 60 s, and it notes that these were not read from any record. And it says that the band holds no sensor lag, so the probability speaks only for the structure drawn, and that the figures are for 20 kW and not for the plant's lower loads. It does not choose a tuning. The decay figure rests on the priors and reads an envelope of the pair, not the settling of a load step.

### 8.5.3 Common mistakes

Table 8.13 lists five mistakes this chapter's plant makes easy. Each row names the mistake, its form on this hall and the correction.

## 8.6 The next measurement (step 7)

The band is almost all a capacity that a steady window cannot read, so the candidates are dynamic tests, and the ranking of Chapter 6 does not carry over.

Table 8.12: The report at 20 kW, $K_i/K_p = 0.02$, fit + dynamic unless marked. State route first, draws (D) beside. Draws: seed 8, $N = 200\,000$, effective size 119 928. The truth's dynamic entries sit at the prior's center.

| Control engineer's question | Answer | Band | Stated truth |
|---|---|---|---|
| What does the band do to the critical gain? | $K_c = 1.0129$; rated 0.8223 | sd 0.0949 (D 0.0983); fit only 0.0145 (D 0.0152) | 1.0136 |
| How likely is this tuning to be stable? | at 0.41: >0.9999 (D >0.9999); at 0.80: 0.9876 (D 0.9918); at 1.00: 0.5541 (D 0.5734) | se of D 0.0002; 0.0014. No draw of 199 999 at or under 0.4112 | stable to 1.0136 |
| What is the largest gain that holds 0.95? | $K_{95} = 0.8568$ (D 0.8693); 30 kW: 1.1868 (D 1.2142); 45 kW: 1.7298 (D 1.7864) | se of D 0.0005; 0.0007; 0.0011. Over the priors of Table 8.10, D 0.9605 to 0.7323 | 1.0136; 1.4405; 2.1871 |
| And if it must also decay? | an envelope of the pair: at 0.41, 0.5263 (D 0.5309); largest gain at 0.95: 0.1992 (D 0.1858) | se of D 0.0014. Requirement 7.5 to 20 minutes, D 0.1174 to 0.9954; 30 to 5 kW, D 0.9520 to 0.0000; 5 s lag, D 0.0453 | holds at 0.41, pair 0.8581; $K_d = 0.4219$ |
| Where does the band come from? | $C_c$ 0.008789 of 0.009009; fit 0.000211 | unseen share 97.7 % (D fit's share 2.4 %); 82.2 % at log-sd 0.05 | not a test of the shares |
| What does instability look like with stops? | at 1.02 $K_c$: $T_c$ p-p 0.90 K, period 5.2 steps, travel about 12 per hour, valve off its stops | a simulation at the truth, no band | the row is the truth |
| Which test would narrow it most? | a set-point step of +0.5 K at $K_p = 0.41$ with a permanent 5 s logger: sd $K_c$ 0.0481 if the noise is white at 5 s; 0.0791 at $\phi = 0.9$ | preposterior; fresh records, white noise, rmse 0.0468 (se 0.0023) | – |

Table 8.13: Five mistakes in reading stability under a band.

| Mistake | On this hall | Correction |
|---|---|---|
| Reading the radius as the decay | radius 0.9445 at 0.10, 0.9492 at 0.41, while the pair goes from 0.7995 to 0.8588 | read the mode that reaches the quantity; put the requirement on the pair |
| Taking a tight fit as a tight stability answer | sd 0.0145 fit only, 0.0949 with what the window cannot see | list what the map needs that no row reads, and give it a prior |
| Comparing fit-only and full as methods | $K_{95}$ 0.9919 against 0.8693; radius 1.0137 at $C_c$ over by 15 % | they are different questions; say which entries were pinned |
| Trusting the first order on a variance term | $\tau_a$ at 0.000003 at 20 kW, with a minimum in $K_c$ beside the center | check by draws where a gradient vanishes |
| Quoting $K_c$ without its line | 1.5588, 1.0129, 0.2224 at $K_i/K_p = 0.01$, 0.02, 0.04 | state the ratio with the critical gain |

### 8.6.1 The preposterior of a test

A candidate's rows give the preposterior of equation (6.23). For a test the rows are the readings over its 40 steps, and their Jacobian is taken by sensitivities at the fit. The form is linear-normal. The logged command is replayed, so a reading given earlier commands carries no feedback term; identification under feedback beyond this one sentence is Chapter 24's. The covariance does not depend on the readings' values. There are seven entries here, $(g, g_w, b, \ell_c, \ell_a, \ell_h)$ and the test's own load offset d$Q$, with a prior spread of $1/\sqrt{40}$ kW from the test's meter. $K_c$ and $K_d$ are carried by their gradients at the fit, so a candidate compares with the state route's present figures, sd $K_c$ 0.0949, $K_{95}$ 0.8568 and $K_d$ at 0.95 of 0.1992, and not with the draws. A valve bump is the one open-loop node of the chapter: it pins the actuator's command.

### 8.6.2 The candidates

Table 8.14 lists the candidates, in no order of merit. Each row is one; the columns give the spreads of $\ln C_c$ and $\ln \tau_a$ after it, then sd $K_c$, $K_{95}$ and the decay's largest gain at 0.95, the last three by the state route. Figure 8.6 draws sd $K_c$ for each as a bar, with the present value dotted.

Table 8.14: The candidates by the linear-normal preposterior, 40 steps. Today: 0.1500, 0.1500, 0.0949, 0.8568, 0.1992.

| Candidate | sd $\ln C_c$ [–] | sd $\ln \tau_a$ [–] | sd $K_c$ [1/K] | $K_{95}$ [1/K] | $K_d$ at 0.95 [1/K] |
|---|---|---|---|---|---|
| Reference discharge sensor, 0.05 K | 0.1500 | 0.1500 | 0.0949 | 0.8568 | 0.1993 |
| Load meter, 0.2 kW | 0.1500 | 0.1500 | 0.0949 | 0.8568 | 0.1992 |
| Record's load step, 30 to 45 kW, $K_p = 0.10$ | 0.1494 | 0.0936 | 0.0940 | 0.8583 | 0.2249 |
| Set-point +1 K, 30 kW, $K_p = 0.10$ | 0.1344 | 0.0555 | 0.0848 | 0.8735 | 0.2627 |
| Valve bump +0.20, 20 kW, open, 30 s log | 0.1410 | 0.0614 | 0.0895 | 0.8657 | 0.2486 |
| Set-point +0.5 K, 20 kW, $K_p = 0.41$ | 0.1195 | 0.0513 | 0.0760 | 0.8879 | 0.2812 |
| Set-point +0.25 K, 20 kW, $K_p = 0.80$ | 0.1140 | 0.0415 | 0.0725 | 0.8937 | 0.2860 |
| Valve bump +0.20, 20 kW, open, 5 s logger | 0.1174 | 0.0275 | 0.0755 | 0.8888 | 0.2817 |
| **Set-point +0.5 K, $K_p = 0.41$, permanent 5 s logger** | **0.0738** | **0.0235** | **0.0481** | **0.9339** | **0.3357** |
| The same, AR(1) noise on the 5 s rows, $\phi = 0.5$ | 0.1027 | 0.0383 | 0.0657 | 0.9049 | – |
| The same, $\phi = 0.7$ | 0.1160 | 0.0484 | 0.0738 | 0.8914 | – |
| The same, $\phi = 0.9$ | 0.1245 | 0.0570 | 0.0791 | 0.8829 | – |

Chapter 6's ranking reverses. Its best sensor removes nothing from this answer: a reference on the discharge sensor leaves sd $K_c$ at 0.0949, since the fit's part of the band was already small. The load meter does the same. Only a test that moves the plant reaches $C_c$ and $\tau_a$. The size of the step and the logger count about alike: on the 30 s log at $K_p = 0.41$, doubling the step from +0.25 to +0.5 K takes sd $K_c$ from 0.0900 to 0.0760, and the logger at +0.25 K takes it to 0.0754. The coil's time constant is 21.9 s at 30 kW at the fit, shorter than the 30 s period, so a log at 30 s steps over part of what is to be read. At 20 kW the bump and the closed steps are matched at about 0.20 in command, so the open and closed rows are comparable there, like with like: on the 30 s log the bump gives 0.0895 against the closed step's 0.0760, and on the logger 0.0755 against 0.0481. The closed step beats the bump on either log.

**The candidates tested on fresh records.** The preposterior is a linear-normal prediction. Two candidates were run on fresh records: seeds 8000 to 8199, with the truth drawn from the prior block, the record's load wander and the sensors' noise, and the MAP on the logged

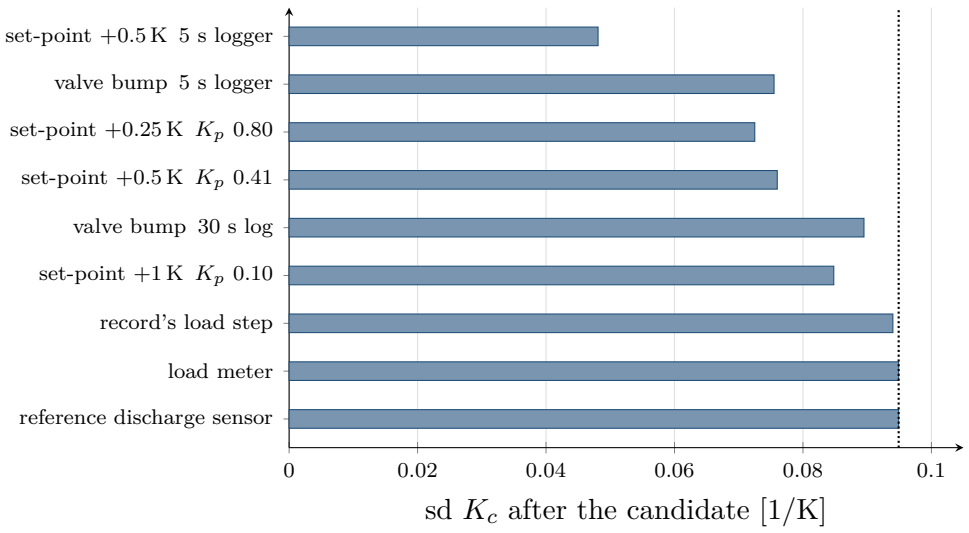



Figure 8.6: The spread of the critical gain after each candidate of Table 8.14, by the preposterior (6.23) with $K_c$ by its gradient at the fit. Dotted: the state route's present 0.0949.

command. The rmse of $K_c$ at 20 kW is 0.0468 (se 0.0023) against the prediction of 0.0481 for the logger, with a mean error of $-0.0029$; and 0.0757 (se 0.0038) against 0.0755 for the bump, with a mean error of $-0.0059$. Unlike the truth row, this test does draw the capacities. It uses the same noise, model and prior as the prediction, so it tests only the linearization. The logger's advantage rests on another assumption, white sensor noise at 5 s. With AR(1) noise on the 5 s rows, at $\phi = 0$, 0.5, 0.7 and 0.9, sd $K_c$ is 0.0481, 0.0657, 0.0738 and 0.0791, against 0.0760 for the 30 s log. Noise correlated that strongly removes the advantage, and these fresh records cannot show it.

### 8.6.3 The recommendation

The next measurement is the permanent 5 s logger, used with a set-point step of +0.5 K at the sheet's $K_p = 0.41$. It takes sd $\ln C_c$ from 0.1500 to 0.0738, sd $K_c$ from 0.0949 to 0.0481 and $K_{95}$ from 0.8568 to 0.9339 by the state route, and it is the permanent instrument, so it reads every later step as well. Those figures hold if the sensor's noise is white at 5 s; at $\phi = 0.9$ the logger gives 0.0791 and no longer beats the 30 s log. The step at $K_p = 0.80$ is a stronger test on the 30 s log, 0.0725 against 0.0760, but it carries Pr(stable) = 0.9918 by the draws, and a miss is the bounded cycle of Section 8.3.6. The dynamic fit that actually uses such a record is work that Chapter 24 owes, as closed-loop identification; the record's own load step is weak, as the table shows, and its set-point step reaches the stop.

### 8.6.4 What the analysis showed

The sheet took one critical gain from a coil that had been overturned and divided by two; the analysis finds that for stability alone this was safe with room at that load and with no sensor lag, on a weaker coil that happens to raise the critical gain, and that a chosen envelope of the pair is about even odds under stated priors. The envelope is not the speed the gain was raised for: raising the gain buys a little on load steps, set by the hall, and costs on set-point steps. A tight posterior and a wide answer are not in tension: the fit reads the steady product and the path, and the critical gain also needs a capacity it never saw. The radius was the wrong measure of what the engineer wanted, since the hall's own slow mode sets it, and that mode barely reaches the discharge error after a kick on $T_c$, while a load step fills it. Past the critical gain the loop on this example does not run away: it settles into a cycle of about one kelvin,

which is valve travel and not a threat to the hall. What would improve the answer is a record that moves, read faster than the coil.

### 8.6.5 Limits of the model

**Structure is not in the band, and this one is large.** A first-order discharge sensor lag, absent from the plant and so from every probability above, moves $K_c$ at the fit to 0.9298, 0.7081, 0.5599 and 0.4160 at 1, 5, 10 and 20 s. A 5 s lag lies below $K_{95}$. Table 8.15 carries the lag through the draws. Each row is a lag, and the columns give $K_c$ at the fit, $K_{95}$ and the two probabilities at the sheet's gain.

Table 8.15: A discharge sensor lag on the draws, fit + dynamic, 20 kW. The first row is the chapter's answer.

| Lag [s] | $K_c$, fit [1/K] | $K_{95}$ [1/K] | Pr(decay at 0.41) [–] | Pr(stable at 0.41) [–] |
|---|---|---|---|---|
| 0 | 1.0129 | 0.8693 | 0.5309 | >0.9999 |
| 1 | 0.9298 | 0.8031 | 0.4070 | >0.9999 |
| 5 | 0.7081 | 0.6225 | 0.0453 | >0.9999 |
| 10 | 0.5599 | 0.4967 | 0.0000 | 0.9999 |
| 20 | 0.4160 | 0.3694 | 0.0000 | 0.6052 |

So the headline's even odds do not survive a lag of a few seconds, and the room in stability is used up near 20 s. The sheet's factor of 2 covers things no prior was given, up to about there; the 0.95 does not. These are two different guarantees, and the posterior does not replace the margin.

**The priors on the dynamic entries.** They are stated and not read, and their centers are the plant's own numbers. The 97.7 % does not stand without them.

**The plant's form.** The valve is linear in $u$, the coil one lump. $K_i/K_p$ is fixed. Loads below 20 kW carry a lower $K_c$, and 20 kW is the sheet's load, not the plant's. At $K_p = 0.41$, by the draws, the decay holds with 0.9520, 0.5309, 0.1606, 0.0104 and 0.0000 at 30, 20, 15, 10 and 5 kW; at 5 kW, $K_{95}$ is 0.4198 and stability 0.9651; and the truth's $K_c$ at 2 kW is 0.398. The window was at about 30 kW, so 20 kW is an extrapolation on the model.

**The event.** Stability here is of the unsaturated map only. The cycle was found by simulation at the truth, with the load fixed during it, and is given no probability. What the stop does as a regime is Chapter 15's.

**The sampler.** The critical gain is largely a property of the sampling. With the ratio held, $K_c$ at the fit and 20 kW is 1.0129, 1.8201, 4.8734 and 23.1937 at controller periods of 30, 15, 5 and 1 s. The plant's controller is an ideal sampler at 30 s: its rate limit, deadband and latency are unstated, and each would move the figure.

**The requirement and the percentages.** The 10-minute requirement is a choice. Percentages are this example's.

**The constructed record.** The record is constructed from a stated truth, so the truth's agreement tests the method on its own model, not the model on a real hall. Chapter 9 asks what a moving record can say about the state.

## 8.7 What to retain

**Definition 8.1** (Spectral radius)**.** For a one-period map $\mathbf{M}$ on the deviations of a sampled loop, $\rho(\mathbf{M}) = \max_i |\lambda_i|$. The deviations die from every start exactly when $\rho(\mathbf{M}) < 1$. The radius answers whether all entries settle; it is not the rate at which a chosen quantity settles, which depends on the residue of each mode on it.

Three conclusions on this hall go with the method. Stability under a band is a threshold question, on a scalar root when the check of Section 8.2.4 says so. What a steady window reads and what the map needs differ, and a variance budget shows the difference; here the part not read is most of the band, on stated priors. And what a word such as stable buys has to be read from the plant: here, at the sheet's load and with no sensor lag, the sheet's gain is stable with room and about even on a chosen envelope of the pair, which is not the settling of a load step; and past the limit the cost is valve travel.

Table 8.16 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 8.16: The terms of the chapter and where each is defined. Units for the closed-loop hall.

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| Modes | eq. (8.3) | – | a state as a sum of eigenvectors | 8.1.1 |
| Spectral radius | eq. (8.4) | – | largest modulus of the map | 8.1.1 |
| One-period map | eq. (8.6) | mixed | deviations over one step | 8.1.2 |
| Prior on a capacity | $\ell_c/0.15$ | – | what no steady row reads | 8.1.2 |
| State route, draws route | Section 8.2.3 | – | two readings of one posterior | 8.2.3 |
| Critical gain | $\rho(\mathbf{M}(K_c;\mathbf{z})) = 1$ on a stated line | 1/K | one scalar root | 8.2.4 |
| Probability of stability | eq. (8.13) | – | a threshold on $K_c$ | 8.3.2 |
| $K_{95}$ | $\Pr(K_c > K_{95} \mid \mathbf{y}) = 0.95$ | 1/K | the gain holding the level | 8.3.3 |
| The controller's pair | hall mode set aside, $\lvert\lambda_{\text{pair}}\rvert$ | – | the mode a gain moves | 8.2.4 |
| Decay requirement | $\lvert\lambda_{\text{pair}}\rvert \le 0.05^{1/20}$ | – | a chosen envelope of the pair | 8.3.4 |
| Unseen share | eq. (4.54) on the last three entries | % | what the window cannot read | 8.3.5 |
| Limit cycle | simulation at the truth | K | size set by the first clip | 8.3.6 |
| Preposterior of a test | eq. (6.23) | mixed | the band after a test's rows | 8.6.1 |

## 8.8 Exercises

**Exercise 8.1.** *Formulation.* A sampled PI controller holds the level of a tank by moving a pump's speed, at a stated $K_i/K_p$. Write the event and the scenario node for the question "is this tuning stable, given the record?", and name the entry a steady window cannot read. (The node pins the set point and the inflow and leaves the pump's deviation free, closed loop. The event is $\rho(\mathbf{M}(K_p;\mathbf{z})) < 1$ on the one-period map of the level, the pump's lag and the integrator, under equation (6.1) with $f = \rho$ and $f^* = 1$. The tank's area divides the level's rate and appears in no steady row, and neither does the pump's time constant, so each takes a prior row.)

**Exercise 8.2.** *Derivation.* Show that a steady window reads $\ln(Q/u)$, and that $K_c$ moves as $1/W$ with the rest held. (The coil's row at steady state gives $W(T_c - T_w)u = Q$, so $\ln[W(T_c - T_w)] = \ln(Q/u)$, both read: its spread is 0.00596. In $\mathbf{A}$ of equation (8.5) the entry $Wu$ equals $Q/(T_c - T_w)$, which holds no $W$; $W$ appears only in the valve's column as the product.

Scaling $W$ by $s$ and scaling $\delta u$, $\delta I$ and $K_p$ by $1/s$ leaves the map's eigenvalues unchanged, so $K_c \propto 1/W$ and the gradient in $g_w$ is $-1$ in the log.)

**Exercise 8.3.** *Derivation.* Say why the offset and the coil cancel in $K_c$. (Their gradient entries in the log have opposite signs, $+0.4413$ on $b$ and $-1$ on $g_w$, and $\mathrm{corr}(g_w, b) = 0.9924$, so $2\,\partial_{g_w} K_c\,\partial_b K_c\,\Sigma_{g_w b}$ is negative, $-0.008268$. The split of Section 8.3.5 says why: the gradient is $-1$ times the line of the product the valve reading fixes and $-0.1750$ times the line the hall reading fixes, with 0.0127 left on $b$, and the fit's total falls to 0.000211.)

**Exercise 8.4.** *Derivation.* Say why the hall mode leaves the discharge error. (The eigenvector at $K_p = 0.41$ with the $T$ entry set to one is $(1,\ -0.0235,\ 0.1195,\ 0.1137)$: the loop holds $T_c$ while the hall air relaxes. In the sum (8.3), a unit kick on $T_c$ puts into this mode a residue of 0.0008 K on $T_c$ against the pair's 0.5494 K, so after that kick the error's decay is the pair's, though the radius is the hall's 0.9492. A load step moves the hall air itself and does fill this mode, as Table 8.5 shows.)

**Exercise 8.5.** *Application.* Find $K_{95}$ at 30 kW, fit + dynamic. (State route: $K_c = 1.4130$, the budget's total 0.018916, so a spread of 0.1375 and $K_{95} = 1.4130 - 1.6449 \times 0.1375 = 1.1868$. Draws: the weighted 5 % quantile of the root over the $t_5$ draws with the dynamic entries drawn, se 0.0007, which is 1.2142.)

**Exercise 8.6.** *Application.* Find Pr(decay) at $K_p = 0.41$ with a 15-minute requirement, fit + dynamic. (Thirty steps give $|\lambda_{\mathrm{pair}}| \le 0.05^{1/30} = 0.9050$. At the MAP state $K_d = 0.6051$. On each draw solve for $K_d$ on the map's pair, set $K_d = 0$ where the requirement fails already at 0.10, and sum the weights with $K_d > 0.41$: 0.9502.)

# Part III

# The Belief

# Chapter 9

# The Kalman Filter as a Time-Stacked Fit

## 9.1 A filter block on the hall sensor

The commissioning engineer has the vendor's sheet open at a filter block on the hall sensor. Its gain is 0.05 per step and it is offered as a device, separate from any model: the reading's noise falls from 0.500 K to 0.0801 K, with a lag of 570 s. The commissioning sheet's fit of the hall stays where it was, in another tab. The load has just stepped to 45 kW. Four steps later the block's trend reads 22.911°C against the 22.875°C of the step itself, and the hall looks as though it has barely moved.

The engineer would conclude that the block is one more piece of equipment, with a gain to be bought and a band of 0.0801 K on the hall. The sheet does not say whether that device has anything to do with the fit in the other tab, or what its number means while the hall is moving. The question of this chapter is

*whether the filter is a device with a gain to be bought or a fit that has already chosen its rows, and what its band means through a step of the load.*

**Jargon: time-stacked state, Kalman filter, predict and update, innovation, smoother, steady gain, virtual reading, one-pass extended filter, iterated stacked fit, false pin**

The *time-stacked state* is the column of one state per controller step, all of them entries of one fit. The *Kalman filter* is a recursion that carries a mean and a variance from step to step. Its *predict* carries the belief one step through the dynamics row, and its *update* adds the new reading's row. The *innovation* is the part of a reading the prediction did not expect. The *smoother* gives the belief on an earlier step from the readings after it as well. The *steady gain* is the value the recursion's gain settles to. A *virtual reading* is a quantity not read at that moment, or not read well, reported from the fit with its band. The *one-pass extended filter* runs the recursion on rows that are not linear, linearizing each row once at the mean of the moment, and the *iterated stacked fit* solves all the rows together and relinearizes every one at the final state. A *false pin* is a quantity held fixed in the rows that is really uncertain, or held at a value it does not have.

**Key idea**

Stack the state of every step into one fit and eliminate the entries one at a time, oldest first. Eliminating an entry is the predict, and adding the next reading's row is the update. On linear rows with normal errors the recursion is therefore the last step's marginal of the stacked fit, by algebra. A gain on a sheet is a choice of rows, and a band is only as wide as what those rows hold fixed.

### 9.1.1 The Fit line of the control problem

Chapter 7 wrote what the policy may see as one line of the problem (7.1), $(\hat{\mathbf{z}}_k, \mathbf{\Sigma}_k) = \mathrm{Fit}(\hat{\mathbf{z}}_{k-1}, \mathbf{\Sigma}_{k-1}, \mathbf{u}_{k-1}, \mathbf{y}_k)$, and Section 7.2.1 used one step of it ahead of its derivation. This chapter derives that line on the hall's air node, with one difference stated: Fit takes $\mathbf{u}_{k-1}$, this filter's input is the metered $y_{Q,k-1}$, and $u$ has no row here. It also fixes the notation the rest of the book keeps. The state at controller step $k$, 30 s apart, is $\mathbf{z}_k$. The stack from step $K_0$ is $\mathbf{z}_{K_0:k} = (\mathbf{z}_{K_0}^\mathsf{T}, \dots, \mathbf{z}_k^\mathsf{T})^\mathsf{T}$ and the readings over those steps are $\mathbf{y}_{K_0:k}$. The belief on step $k$ from the readings through step $j$ has the moments

$$\hat{\mathbf{z}}_{k|j} = \mathbb{E}[\mathbf{z}_k \mid \mathbf{y}_{K_0:j}], \qquad \mathbf{\Sigma}_{k|j} = \mathrm{Cov}(\mathbf{z}_k \mid \mathbf{y}_{K_0:j}). \tag{9.1}$$

The predicted belief is $k \mid k-1$. The filter's belief is $k \mid k$, which is Chapter 7's $(\hat{\mathbf{z}}_k, \mathbf{\Sigma}_k)$. The smoother's is $k \mid N$, with $N$ the window's last step. The gain is $\mathbf{\Upsilon}_k$, the symbol of equation (3.29), clear of a controller's gains. Notation follows dimension: while the state of one step has the single entry $T_k$, these are the numbers $\hat{T}_{k|j}$, $P_{k|j}$ and $\upsilon_k$, and the bold forms return when $g$ joins the state.

### 9.1.2 The hall air over a window (step 0)

The plant is the closed-loop hall's air node, over steps 60 to 120 of the record, both included: 61 readings over 30 minutes at $r = 18°\mathrm{C}$, with the load stepping from 30 to 45 kW at step 80. The node's balance is $C_h \dot{T} = Q - G(T - T_c)$, with $C_h = 3600\,\mathrm{kJ/K}$. With $T_c$ and $Q$ held over the step of $\Delta t = 30\,\mathrm{s}$, it integrates to

$$T_{k+1} = a\,T_k + (1-a)\,T_c + \phi\,Q_k + w_k, \qquad a = \mathrm{e}^{-G\,\Delta t/C_h}, \quad \phi = \frac{1-a}{G}, \tag{9.2}$$

and $w_k$ holds what the held values leave out. The two readings are

$$y_{T,k} = T_k + \varepsilon_{T,k}, \qquad y_{Q,k} = Q_k + \varepsilon_{Q,k}, \tag{9.3}$$

with spreads of 0.50 K and 1.00 kW. The record is closed loop. The set point is pinned and the valve is free, but the coil and the valve have no row on this node. They reach it only through $T_c$, and the discharge reading has no row here. The free entries are the $T_k$ and the $Q_k$; $g$ is pinned until Section 9.2.4 frees it. At Chapter 5's fit, $G = 6.1827\,\mathrm{kW/K}$, so $a = 0.949783$, $\phi = 0.008122\,\mathrm{K/kW}$ and $C_h/G = 582.3\,\mathrm{s}$.

Figure 9.1 draws the window. The green boxes along the middle are the hall air at successive steps, the black squares the dynamics rows between neighbors, the gray box above the pinned supply air that enters every dynamics row, and the dashed arrows the two readings: the hall sensor on each $T_k$ and the load meter on each dynamics row. The shaded region on the left holds what Section 9.2.2 eliminates first.

**The modeling choice in the input row.** The row needs a value of $T_c$, and the hall has no reading of it that the loop has not consumed. Chapter 6 established that the controller holds the reading and not the air, so that $T_c = r - b$, with $b = +0.30\,\mathrm{K}$ in the record's stated truth

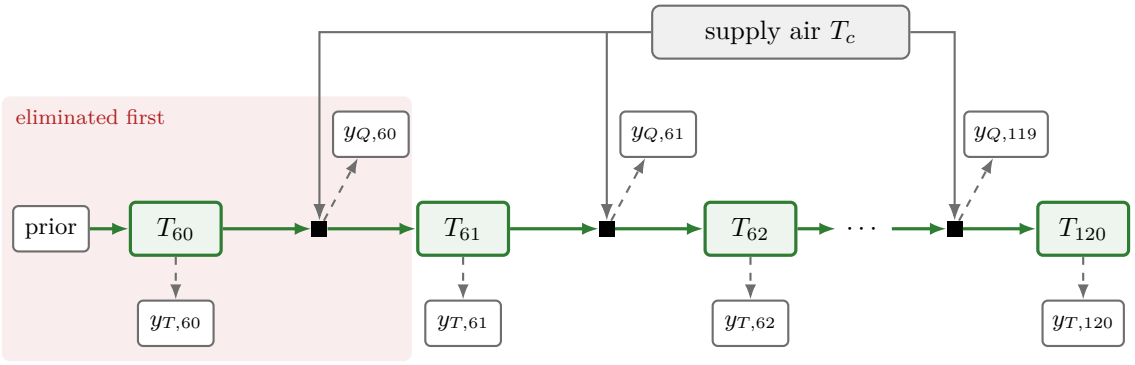



Figure 9.1: The hall air as a chain of steps 60 to 120. Green: the free entries $T_k$. Black squares: the dynamics rows (9.2), each with its free load $Q_k$. Gray: the supply air, pinned in these rows. Dashed: the readings (9.3). Shaded: the entry $T_{60}$ with the rows that touch it.

(Section 6.4.4). The input row must therefore say which $T_c$ it uses, and this chapter keeps one rule. **The headline rows read $T_c = r$, Chapter 5's pair, with $g = 0.02999$.** Chapter 6's pair, $T_c = r - 0.0571$ with $g = 0.01833$ and $G = 6.1110\,\mathrm{kW/K}$, stands beside it wherever the choice is compared. The choice buys rows that use only what a fit knows, and it fails by the offset it leaves out, which Section 9.3.5 measures.

**The vendor's block.** The block keeps one number and moves it toward each reading, $\bar{w}_k = \bar{w}_{k-1} + 0.05\,(y_{T,k} - \bar{w}_{k-1})$. With $\alpha = 0.05$ and a hall that does not move, its noise is $0.50\sqrt{\alpha/(2-\alpha)} = 0.0801\,\mathrm{K}$ and its lag $(1-\alpha)/\alpha = 19.0$ steps. The block and its gain are constructed for this chapter. No vendor's sheet is being quoted.

### 9.1.3 The control engineer's questions (step 1)

The engineer asks six questions of the window. Table 9.1 lists them. The first column is the question in the engineer's words, the second the general form of its category, the third its mathematical form on this hall, the fourth its category, the fifth the computation and the last the subsection. All six read moments of a posterior, of an entry or of a derived quantity.

### 9.1.4 The evidence as rows (step 2)

**The stated choices.** Two rows need values no reading supplies. The prior on the first entry is $T_{60} \sim \mathcal{N}(22.8617,\ 0.50^2)$ in °C, Chapter 5's node with one reading's spread. It buys a start for the chain, and the window's overlap with that fit is a limit stated in Section 9.6.4. The noise of the dynamics row is $\mathrm{sd}(w_k) = 0.02\,\mathrm{K}$ per step. This is a choice, made before any comparison with the truth. It equals $0.40\,\mathrm{K}$ of $T_c$ or $2.46\,\mathrm{kW}$ of unmetered load over one step, and it buys a chain whose neighbors are tied without being pinned to each other. It fails if the row's real error persists from step to step, since white noise cannot represent that, and Section 9.3.5 tests how far it does. In the record's stated truth, unknown to the fit, the row's one-step residual at the truth's own $g$ has a spread of $0.00300\,\mathrm{K}$ with $T_c$ at its window mean, and $0.00035\,\mathrm{K}$ with its own $T_c$ at each step. At Chapter 5's pair, which the headline rows read, the same residual has a mean of $+0.0097\,\mathrm{K}$ per step after the load step, against $+0.0017\,\mathrm{K}$ before it: an error that persists.

**Eliminating an entry.** Every step of what follows uses one operation, written here once. Suppose the rows are linear and their errors normal, with information $\mathbf{\Lambda}$ and right side $\boldsymbol{\eta}$ partitioned into an entry $x$ to remove and the rest $\mathbf{v}$. The objective is $\frac{1}{2}\mathbf{z}^\mathsf{T}\mathbf{\Lambda}\mathbf{z} - \boldsymbol{\eta}^\mathsf{T}\mathbf{z}$ plus a

Table 9.1: The questions asked of the hall's window. $\mathbf{y}_{60:k}$ holds the readings through step $k$, with the prior rows.

| Control engineer's question | General form | Mathematical form | Category | Computation | Answered in |
|---|---|---|---|---|---|
| Is the filter a separate device? | moments, stacked, (2.10) | $p(T_k \mid \mathbf{y}_{60:k})$ against the recursion | estimation | posterior | 9.3.1 |
| What is the hall at the last step? | moments, (2.10) | $\hat{T}_{120\mid120}$, $P_{120\mid120}$ | estimation | posterior | 9.3.2 |
| What do later readings say about earlier steps? | moments, (2.10) | $\hat{T}_{k\mid120}$ against $\hat{T}_{k\mid k}$ | estimation | posterior | 9.3.3 |
| How fast does the band fall, and to what gain? | moments of a derived quantity, (2.10) | $\upsilon_k$, $P_{k\mid k}$, the steady limit | estimation of a derived quantity | posterior | 9.3.4 |
| What did the hall do through the load step? | moments, (2.10) | $T_k$ as a virtual reading, $k = 81, \ldots, 120$ | estimation | posterior | 9.3.5 |
| With $g$ free, does the recursion still give the marginal? | moments of a derived quantity, (2.10) | $p(g, T_{120} \mid \mathbf{y}_{60:120})$, three routes | estimation of a derived quantity | posterior | 9.3.6 |

constant. Completing the square in $x$ leaves a normal density in $x$, whose integral does not depend on $\mathbf{v}$, times

$$\exp\Big[-\tfrac{1}{2}\mathbf{v}^{\mathsf{T}}(\mathbf{\Lambda}_{vv} - \mathbf{\Lambda}_{vx}\Lambda_{xx}^{-1}\mathbf{\Lambda}_{xv})\mathbf{v} + (\boldsymbol{\eta}_v - \mathbf{\Lambda}_{vx}\Lambda_{xx}^{-1}\eta_x)^{\mathsf{T}}\mathbf{v}\Big]. \tag{9.4}$$

The matrix in the first bracket is a Schur complement. It is the block inversion behind equation (3.29), $\mathbf{P}_{zz} = \mathbf{S}^{-1}$, with covariance and information exchanged. The marginal of $\mathbf{v}$ is thus normal, with that information and that right side, and nothing has been approximated.

**The load.** $Q_k$ appears in two rows only: the dynamics row and its meter. With $d_k = T_{k+1} - aT_k - (1-a)T_c$ and the meter's variance $\sigma_Q^2$, equation (9.4) with $x = Q_k$ gives

$$\min_{Q_k}\left[\frac{(d_k - \phi Q_k)^2}{\sigma_w^2} + \frac{(Q_k - y_{Q,k})^2}{\sigma_Q^2}\right] = \frac{(d_k - \phi\, y_{Q,k})^2}{\sigma_w^2 + \phi^2\sigma_Q^2}, \tag{9.5}$$

so the load leaves a known input $\phi\, y_{Q,k}$ and a variance

$$q = \sigma_w^2 + (\phi\sigma_Q)^2 = 4.6597 \times 10^{-4}\,\mathrm{K}^2, \qquad \sqrt{q} = 0.02159\,\mathrm{K}. \tag{9.6}$$

Reading the meter as a known input is therefore no extra assumption on these rows: it is the marginal of the rows with the load kept. It does rest on each $Q_k$ being flat apart from its meter. The record's load has a part correlated from step to step, and no row here carries it.

**The rows of the stack.** With $c_k = (1-a)T_c + \phi\, y_{Q,k}$ the input and $R = 0.25\,\mathrm{K}^2$ the hall sensor's variance, the rows on the state $\mathbf{z}_{60:120} = (T_{60}, \ldots, T_{120})^{\mathsf{T}}$, of 61 entries, are one prior row, 61 reading rows and 60 dynamics rows,

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T_{60} - 22.8617)/0.50 \\ (T_k - y_{T,k})/\sqrt{R} \\ (T_{k+1} - a\,T_k - c_k)/\sqrt{q} \end{bmatrix}, \qquad \mathbf{J} = \begin{bmatrix} * & 0 & 0 & \cdots \\ * & 0 & 0 & \cdots \\ * & * & 0 & \cdots \\ 0 & * & 0 & \cdots \\ 0 & * & * & \cdots \\ \vdots & & & \ddots \end{bmatrix}, \tag{9.7}$$

with $\mathcal{L} = \frac{1}{2}\mathbf{r}^\mathsf{T}\mathbf{r}$. The Jacobian's pattern, in the columns $T_{60}, T_{61}, T_{62}, \ldots$ and the rows prior, $y_{T,60}$, dynamics 60, $y_{T,61}$, dynamics 61, and so on, is a property of the graph: a reading touches its own entry, and a dynamics row touches two neighbors. The information matrix of these normalized rows is $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$; with unnormalized rows it is $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$. Its entries are

$$\Lambda_{kk} = \frac{\mathbf{1}[k = 60]}{0.50^2} + \frac{1}{R} + \frac{\mathbf{1}[k > 60]}{q} + \frac{\mathbf{1}[k < 120]\, a^2}{q}, \qquad \Lambda_{k,k+1} = -\frac{a}{q}, \tag{9.8}$$

and all others vanish, so 181 of the 3721 elements are nonzero.

Figure 9.2 draws this. On the left each small square is one element of the $61 \times 61$ matrix, blue where it can be nonzero. On the right the first six entries appear at two stages of the elimination, oldest first: gray marks what has been removed and red the one element that each removal changes.

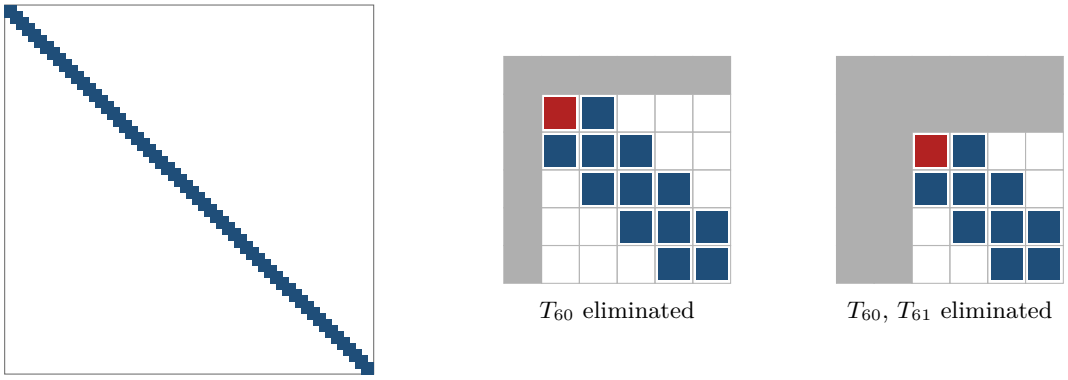



Figure 9.2: The information matrix of the stack, equation (9.8), and its elimination order. Left: the pattern, 181 of 3721 elements nonzero. Right: the first six entries; eliminating the oldest entry by equation (9.4) changes one diagonal element (red) and creates no new nonzero.

The pattern says what the elimination costs: each step touches one element, whatever the window's length, and the pattern stays tridiagonal. The whole of the recursion is in that fact.

**The rows with $g$ free.** Section 9.2.4 frees $g$, on the prior $g \sim \mathcal{N}(0.22314,\ 0.30^2)$, the series' handbook. Chapter 5's fit is not reused as this prior, because its window, steps 40 to 79, overlaps this one. Its node does start the chain, in the prior on $T_{60}$, a reuse Section 9.6.4 states. The load is then kept as 60 entries, because $\phi$ now depends on an entry and $Q_k$ multiplies it.

## 9.2 The posterior of the state (step 3)

All the answers read one of two posteriors, that of the stack with $G$ known and that of the stack with $g$ free. The first is derived twice over, by solving it whole and by eliminating its entries, and that second derivation is the recursion.

### 9.2.1 Two windows by hand

Take the first two entries alone, $\mathbf{z} = (T_{60}, T_{61})^\mathsf{T}$, and the rows that touch them: the prior, $y_{T,60} = 23.03°\mathrm{C}$, the dynamics row with $y_{Q,60} = 29.6\,\mathrm{kW}$, whose input is $c_{60} = 1.14433\,\mathrm{K}$, and

$y_{T,61} = 23.55°\mathrm{C}$. In the form of equation (9.7),

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T_{60} - 22.8617)/0.50 \\ (T_{60} - 23.03)/0.50 \\ (T_{61} - 0.949783\,T_{60} - 1.14433)/0.02159 \\ (T_{61} - 23.55)/0.50 \end{bmatrix}, \qquad \mathbf{J} = \begin{bmatrix} 2.0 & 0 \\ 2.0 & 0 \\ -43.9992 & 46.3255 \\ 0 & 2.0 \end{bmatrix}, \tag{9.9}$$

in $\mathrm{K}^{-1}$. The rows are linear, so the normal equations of Chapter 3, equation (3.25), reach the minimum in one step, and equation (3.26) gives the covariance:

$$\begin{bmatrix} 1943.9263 & -2038.2837 \\ -2038.2837 & 2150.0530 \end{bmatrix} \begin{bmatrix} T_{60} \\ T_{61} \end{bmatrix} = \begin{bmatrix} -2148.9102 \\ 2550.0012 \end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 0.086195 & 0.081714 \\ 0.081714 & 0.077931 \end{bmatrix}. \tag{9.10}$$

The solution is $T_{60} = 23.14593°\mathrm{C}$ and $T_{61} = 23.12872°\mathrm{C}$, with $\mathrm{corr}(T_{60}, T_{61}) = 0.9970$: the dynamics row ties the neighbors much more tightly than either reading fixes them. The entry wanted at the end of these two windows is $T_{61}$, with variance $0.077931\,\mathrm{K}^2$.

### 9.2.2 Eliminating the older entry

Equation (9.4) with $x = T_{60}$ removes the older entry from equation (9.10). Numerically it leaves the information $2150.0530 - 2038.2837^2/1943.9263 = 12.83183\,\mathrm{K}^{-2}$ and the right side 296.78379, so the mean is $23.12872°\mathrm{C}$ and the variance $0.077931\,\mathrm{K}^2$: the entry of the joint solution, without solving for $T_{60}$. Figure 9.1 shows why the operation is small: only three rows touch $T_{60}$.

Now do it in symbols, at any step. Suppose everything before step $k$ has already been eliminated and has left a normal belief on $T_{k-1}$ with mean $\hat{T}_{k-1|k-1}$ and variance $P_{k-1|k-1}$, written $m$ and $P$ for this derivation. The entries $T_{k-1}$ and $T_k$ share the belief's row and the dynamics row,

$$\boldsymbol{\Lambda} = \begin{bmatrix} 1/P + a^2/q & -a/q \\ -a/q & 1/q \end{bmatrix}, \qquad \boldsymbol{\eta} = \begin{bmatrix} m/P - a\,c/q \\ c/q \end{bmatrix}. \tag{9.11}$$

Equation (9.4) with $x = T_{k-1}$ uses $\Lambda_{xx}^{-1} = Pq/(q + a^2P)$, and it leaves on $T_k$ the information

$$\frac{1}{q} - \frac{a^2}{q^2} \cdot \frac{Pq}{q + a^2P} = \frac{1}{q} \cdot \frac{q + a^2P - a^2P}{q + a^2P} = \frac{1}{a^2P + q} \tag{9.12}$$

and the right side

$$\frac{c}{q} + \frac{a}{q} \cdot \frac{Pq}{q + a^2P}\Big(\frac{m}{P} - \frac{a\,c}{q}\Big) = \frac{c\,(q + a^2P) + a\,m\,q - a^2c\,P}{q\,(q + a^2P)} = \frac{a\,m + c}{a^2P + q}. \tag{9.13}$$

Equation (9.12) is the matrix inversion lemma, here in one entry and derived line by line. Its two results, dividing the right side by the information, are the *predict*:

$$\hat{T}_{k|k-1} = a\,\hat{T}_{k-1|k-1} + c_{k-1}, \qquad P_{k|k-1} = a^2P_{k-1|k-1} + q. \tag{9.14}$$

**Eliminating an entry is the predict.** What remains on $T_k$ is one more row, its reading. Adding it is the posterior of equations (3.27) and (3.28) with this prediction as the prior; in the correction form (3.29), with $\mathbf{A} = 1$ and $\mathbf{R} = R$, it reads

$$\frac{1}{P_{k|k}} = \frac{1}{P_{k|k-1}} + \frac{1}{R}, \qquad \upsilon_k = \frac{P_{k|k-1}}{P_{k|k-1} + R}, \qquad \hat{T}_{k|k} = \hat{T}_{k|k-1} + \upsilon_k\,\nu_k, \qquad P_{k|k} = (1 - \upsilon_k)\,P_{k|k-1}, \tag{9.15}$$

with the innovation $\nu_k = y_{T,k} - \hat{T}_{k|k-1}$. This is also the rank-one form (4.74) with $\mathbf{h} = 1$ and $S = P_{k|k-1} + R$, whose $\mathbf{K}$ is $\upsilon_k$, since $P_{k|k-1} - S\upsilon_k^2 = (1 - \upsilon_k)P_{k|k-1}$. **The reading's row added is the update.** The recursion (9.14)–(9.15) is the Kalman filter, and it starts with

the update at $k = 60$, from the prior. Nothing in it was assumed about a filter: it is the fit of equation (9.7), with its oldest entry removed at each step.

Table 9.2 runs the first step by each route. Each row is a quantity and the columns are the three computations. The first two columns solve the rows of equation (9.10); the third uses only the belief at step 60, whose variance is $0.125000\,\mathrm{K}^2$ because the prior and the first reading have equal spreads.

Table 9.2: The two windows by three routes. The information form of the recursion is $1/P_{61|60} + 1/R = 8.83183 + 4.00000$.

| Quantity [unit] | Joint solution | Schur complement | Recursion |
|---|---|---|---|
| Belief at step 60, mean [°C] | – | – | 22.94587 |
| Predicted mean at step 61 [°C] | – | – | 22.93792 |
| Predicted variance at step 61 [K$^2$] | – | – | 0.113227 |
| Gain $\upsilon_{61}$ [–] | – | – | 0.311725 |
| Innovation $\nu_{61}$ [K] | – | – | 0.61208 |
| Information on $T_{61}$ [K$^{-2}$] | – | 12.83183 | 12.83183 |
| Mean of $T_{61}$ [°C] | 23.12872 | 23.12872 | 23.12872 |
| Variance of $T_{61}$ [K$^2$] | 0.077931 | 0.077931 | 0.077931 |

The routes agree because they are one piece of algebra. The computed gaps, $3.2 \times 10^{-13}\,\mathrm{K}$ in the mean, $1.1 \times 10^{-15}\,\mathrm{K}^2$ in the variance and $1.3 \times 10^{-13}$ in the information, are rounding.

### 9.2.3 The whole window

The argument repeats along the chain. Eliminating $T_{60}$ leaves a tridiagonal matrix one entry shorter, whose first entry holds a belief, and so on to the last entry. No step uses anything about the length of the window, so for rows of the form (9.7) with normal errors, the mean and variance the recursion gives at the last step are the last entry's marginal under the stack. Solving the stack through step 120 gives $\hat{T}_{120|120} = 25.039713$°C with variance $0.00407879\,\mathrm{K}^2$, a spread of $0.0639\,\mathrm{K}$. The $61 \times 61$ covariance is too large to display; its last row holds the correlations of $T_{120}$ with the earlier entries, 0.9413 with $T_{119}$, 0.5311 with $T_{110}$ and 0.2751 with $T_{100}$.

### 9.2.4 The stack with $g$ free

Freeing $g$ makes the rows bilinear: $a(g)$ multiplies $T_k$ and $\phi(g)$ multiplies $Q_k$. The state is $\mathbf{z} = (T_{60}, \ldots, T_{120},\ Q_{60}, \ldots, Q_{119},\ g)^{\mathsf{T}}$, of 122 entries, and the 183 rows are the prior on $T_{60}$, the 61 hall readings, the 60 dynamics rows, the 60 meter rows and the prior on $g$,

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T_{60} - 22.8617)/0.50 \\ (T_k - y_{T,k})/0.50 \\ (T_{k+1} - a(g)T_k - (1 - a(g))T_c - \phi(g)Q_k)/0.02 \\ (Q_k - y_{Q,k})/1.00 \\ (g - 0.22314)/0.30 \end{bmatrix}. \tag{9.16}$$

A dynamics row's line in the columns $T_k$, $T_{k+1}$, $Q_k$ and $g$ is

$$\frac{1}{\sigma_w}\Big(-a,\ \ 1,\ \ -\phi,\ \ -a'(g)\,(T_k - T_c) - \phi'(g)\,Q_k\Big), \qquad a' = -a\,\frac{G\,\Delta t}{C_h}, \quad \phi' = -\frac{a'}{G} - \phi, \tag{9.17}$$

and the column of $g$ is the only one every dynamics row shares. At the fit, $a = 0.951938$, $\phi = 0.008131\,\mathrm{K/kW}$, $a' = -0.046888$ and $\phi' = -0.000199$, and the first dynamics row's line is

$(-47.597,\ 50.000,\ -0.4066,\ 11.593)$. Each Gauss–Newton iteration solves $(\mathbf{J}^\mathsf{T}\mathbf{J})\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$, the normal equations (3.25) at the new linearization, from a start at the handbook's $g$ with the $T_k$ run forward through the rows. Table 9.3 gives the iterates: each row is one iteration, and the columns are $g$, the last entry $T_{120}$, the objective and the norm of its gradient. This is the *iterated stacked fit*, named for the route it is.

Table 9.3: Gauss–Newton iterations of the stack with $g$ free, $T_c = r$.

| Iterate [–] | $g$ [–] | $T_{120}$ [°C] | $\mathcal{L}$ [–] | $\lVert\mathbf{J}^\mathsf{T}\mathbf{r}\rVert$ [–] |
|---:|---:|---:|---:|---:|
| 0 | 0.223144 | 23.8623 | 142.84672 | $3.38 \times 10^{1}$ |
| 1 | −0.025967 | 25.2156 | 32.45659 | $1.52 \times 10^{2}$ |
| 2 | −0.014972 | 25.2581 | 30.47747 | $7.92 \times 10^{-1}$ |
| 3 | −0.015013 | 25.2579 | 30.47744 | $4.43 \times 10^{-4}$ |
| 4 | −0.015012 | 25.2579 | 30.47744 | $1.27 \times 10^{-6}$ |
| 5 | −0.015012 | 25.2579 | 30.47744 | $1.84 \times 10^{-9}$ |

The band is the Gauss–Newton covariance, $\mathbf{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ at the fit, with the curvature term $\mathbf{C}$ of equation (4.44) left out; Section 9.3.6 tests that against a route that does not use it. The block of $\mathbf{\Sigma}$ in the entries $(T_{120}, g)$ is

$$\begin{bmatrix} 0.013352 & -0.001890 \\ -0.001890 & 0.000391 \end{bmatrix}, \qquad \operatorname{corr}(T_{120}, g) = -0.8274, \tag{9.18}$$

so $\sqrt{\Sigma_{gg}} = 0.01977$ and the spread on $T_{120}$ is 0.1156 K, wider than the 0.0639 K with $G$ pinned. The correlation reads by sign: a smaller conductance goes with a warmer hall.

## 9.3 The answers (step 4)

Each answer reads one of the two posteriors of Section 9.2. The reported value is the state route, the stack's; the recursions and the grid stand beside it in the same answer, each named.

### 9.3.1 The filter against the stack

**Control engineer's question.** Is the filter a separate device?

**General form.** An estimation question reads moments, equation (2.10), here of one entry of a stack with the other entries integrated out,

$$p(\mathbf{z}_k \mid \mathbf{y}_{K_0:j}) = \int p(\mathbf{z}_{K_0:N} \mid \mathbf{y}_{K_0:j}) \prod_{i \neq k} \mathrm{d}\mathbf{z}_i. \tag{9.19}$$

**Specialized form.** Equation (9.19) with $\mathbf{z}_k = T_k$, $j = k$, $K_0 = 60$ and the stack ending at $k$, on the rows (9.7). Two routes are set side by side: the *stack route*, which solves the normal equations through step $k$ and reads the last entry, and the *filter route*, the recursion (9.14)–(9.15). Their agreement is the algebra of Section 9.2.2, and the numbers test the arithmetic, at a stated tolerance of $10^{-9}$.

**Building blocks.** The stack of Section 9.2.3; the recursion; the load, kept or eliminated by equation (9.5).

**The plant's numbers.** At step 120 both routes give 25.039713°C with variance 0.00407879 $\mathrm{K}^2$, with gaps of $1.3 \times 10^{-13}$ K and $1.5 \times 10^{-17}$ $\mathrm{K}^2$. With the stack re-solved to each of the 61 steps, the largest gaps are $2.9 \times 10^{-13}$ K and $1.1 \times 10^{-15}$ $\mathrm{K}^2$. With the load kept as 60 entries and its meter rows, the last step agrees to $1.5 \times 10^{-13}$ K and $7.4 \times 10^{-17}$ $\mathrm{K}^2$. **On linear rows with normal errors the filter is the marginal at the last step of the stacked fit, and it discards nothing for that entry.** The mean and variance match on linear rows for any noise; "marginal" and "discards nothing" also need the errors normal. What it drops is the joint: the correlations 0.9413, 0.5311 and 0.2751 of Section 9.2.3 are not in its two numbers, and it cannot say what step 120 implies about step 100. So it is no separate device. It is the fit, in an order that spends one element per step.

### 9.3.2 The belief at the last step

**Control engineer's question.** What is the hall at the last step?

**General form.** Moments, equation (2.10), of an entry of the state.

**Specialized form.** Equation (2.10) with $f(\mathbf{z}) = T_{120}$, the marginal (9.19) at $k = j = 120$, under the headline rows, $T_c = r$ and $G$ at Chapter 5's value. The mean is the last entry of $\hat{\mathbf{z}}$ and the variance the last diagonal element of $\boldsymbol{\Sigma}$.

**Building blocks.** $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ of Section 9.2.3; the filter's pair beside them.

**The plant's numbers.** The hall air at step 120 is 25.040°C with a spread of 0.0639 K by the stack, and the filter gives the same to the gaps above. The reading at that step is 25.12°C with a spread of 0.50 K, against which the belief's 0.0639 K is narrow. That width rests on the rows' pinned $G$ and $T_c$.

### 9.3.3 What later readings say about earlier steps

**Control engineer's question.** What do later readings say about earlier steps?

**General form.** Moments, equation (2.10), under the marginal (9.19) with $j = N$ and $k < N$.

**Specialized form.** Equations (2.10) and (9.19) with $\mathbf{z}_k = T_k$ and $N = 120$: on the state route, this is the $k$-th entry of $\hat{\mathbf{z}}$ and the $k$-th diagonal element of $\boldsymbol{\Sigma}$ from the stack of Section 9.2.3. A backward pass over the filter gives the same moments. Given $T_{k+1}$ and the readings through $k$, the entry $T_k$ is normal, by the conditional (3.29) applied to the dynamics row; with $D_k = a\,P_{k|k}/P_{k+1|k}$ its mean is $\hat{T}_{k|k} + D_k(T_{k+1} - \hat{T}_{k+1|k})$ and its variance $P_{k|k} - D_k^2 P_{k+1|k}$. The chain of Figure 9.1 says that the later readings reach $T_k$ only through $T_{k+1}$, so averaging over the belief on $T_{k+1}$, with the total variance rule (3.32), gives the *smoother*:

$$\hat{T}_{k|N} = \hat{T}_{k|k} + D_k(\hat{T}_{k+1|N} - \hat{T}_{k+1|k}), \qquad P_{k|N} = P_{k|k} + D_k^2(P_{k+1|N} - P_{k+1|k}), \tag{9.20}$$

run from $k = N - 1$ down to $k = 60$. A value at step $k$ uses readings that did not exist at step $k$, so a smoothed value is for the log and not for the loop's next move.

**Building blocks.** The filter's predicted and updated pairs, and the stack's diagonal.

Table 9.4: The filter's belief $k \mid k$ against the stack's $k \mid 120$.

| | Filter | | Stack | | | | Error [K] | |
|---|---|---|---|---|---|---|---|---|
| $k$ [–] | Mean [°C] | sd [K] | Mean [°C] | sd [K] | Ratio [–] | Truth [°C] | Filter | Stack |
| 60 | 22.946 | 0.3536 | 22.961 | 0.1591 | 0.450 | 22.882 | +0.064 | +0.078 |
| 70 | 22.838 | 0.1137 | 22.899 | 0.0958 | 0.843 | 22.893 | −0.056 | +0.006 |
| 80 | 22.879 | 0.0774 | 22.914 | 0.0710 | 0.917 | 22.910 | −0.031 | +0.004 |
| 85 | 23.442 | 0.0709 | 23.498 | 0.0659 | 0.930 | 23.499 | −0.057 | −0.001 |
| 90 | 23.891 | 0.0674 | 23.924 | 0.0631 | 0.936 | 23.973 | −0.082 | −0.049 |
| 100 | 24.457 | 0.0647 | 24.507 | 0.0611 | 0.944 | 24.572 | −0.115 | −0.065 |
| 110 | 24.822 | 0.0640 | 24.853 | 0.0613 | 0.957 | 24.909 | −0.087 | −0.055 |
| 115 | 24.955 | 0.0639 | 24.960 | 0.0621 | 0.971 | 25.028 | −0.073 | −0.068 |
| 120 | 25.040 | 0.0639 | 25.040 | 0.0639 | 1.000 | 25.124 | −0.084 | −0.084 |

**The plant's numbers.** The backward pass agrees with the stack's diagonal to $6.7 \times 10^{-13}\,\mathrm{K}$ and $1.1 \times 10^{-16}\,\mathrm{K}^2$. Table 9.4 sets the two beliefs side by side: each row is a step, the columns give the filter's mean and spread, the stack's mean and spread, the ratio of the spreads, the stated truth of the record, and each mean's error, estimate minus truth.

Later readings narrow the first step most, to 0.450 of its filter's spread, because a single early reading had little behind it. From step 80 to step 110 the band narrows only by 4 to 8 %. The larger change is in the mean: the error at step 85 moves from $-0.057$ to $-0.001\,\mathrm{K}$, because the readings after the load step correct the climb the filter was following from behind. Across steps 81 to 120 the RMS error against the truth falls from 0.0915 to 0.0597 K, on this one record. At the last step nothing is later and the two beliefs agree, as Section 9.3.1 requires.

### 9.3.4 The gain and the band

**Control engineer's question.** How fast does the band fall, and to what gain?

**General form.** Moments, equation (2.10), of the entry $T_k$ through the marginal (9.19), with the variance and the gain read as quantities derived from it.

**Specialized form.** Equations (2.10) and (9.19) at $j = k$, whose variance and gain are those of the recursion (9.14)–(9.15). Neither line of the variance contains a reading, so the sequence is fixed by the rows and the prior. Its limit follows from setting $P^- = P_{k|k-1}$ constant in the recursion. With $P = P^-R/(P^- + R)$ and $P^- = a^2P + q$, multiplying through by $P^- + R$ gives

$$(P^-)^2 + \beta\,P^- - qR = 0, \quad \beta = R(1-a^2) - q, \quad P^- = \tfrac{1}{2}\Big(-\beta + \sqrt{\beta^2 + 4qR}\Big), \quad \upsilon = \frac{P^-}{P^- + R}. \tag{9.21}$$

**Building blocks.** $a$ and $q$ of equations (9.2) and (9.6); $R$; the prior's variance.

**The plant's numbers.** With $\beta = 2.401230 \times 10^{-2}\,\mathrm{K}^2$, the limit is $P^- = 0.004138\,\mathrm{K}^2$, a gain of 0.01628 and an updated spread of 0.0638 K. Table 9.5 gives the sequence, each column a step and the last the limit. Figure 9.3 draws it. In the upper panel the horizontal axis is the step and the vertical axis the spread of the belief, with the filter's spread, the stack's spread and the limit dotted. In the lower panel the same axis carries the gain on a logarithmic scale, with the limit and the vendor's 0.05 dotted.

The band is within a factor 1.05 of its limit at step 91, 930 s after the first reading, and within 1.01 at 1290 s. The load step leaves no mark on either curve, since neither contains a

Table 9.5: The gain and the spread of the filter's belief through the window, with the steady limit of equation (9.21).

| $k$ [–] | 60 | 61 | 65 | 70 | 80 | 100 | 120 | limit |
|---|---|---|---|---|---|---|---|---|
| Gain $v_k$ [–] | 0.50000 | 0.31172 | 0.10826 | 0.05169 | 0.02397 | 0.01677 | 0.01632 | 0.01628 |
| $\sqrt{P_{k\|k}}$ [K] | 0.3536 | 0.2792 | 0.1645 | 0.1137 | 0.0774 | 0.0647 | 0.0639 | 0.0638 |

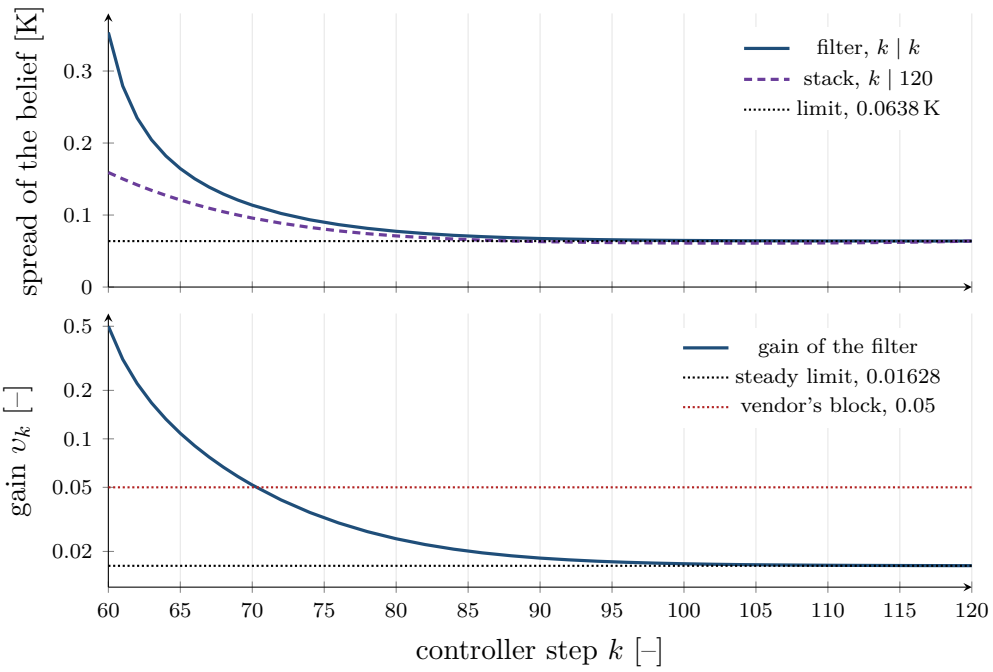



Figure 9.3: The band and the gain of the recursion (9.14)–(9.15) at $\mathrm{sd}(w_k) = 0.02\,\mathrm{K}$. Upper: the spread of the filter's belief and of the stack's. Lower: the filter's gain. The limits are from equation (9.21).

reading: adding 1 K to every reading changes the variances by 0.0. What one reading does to a belief is the question of Chapter 10. In the upper panel the stack's spread is smallest in the middle of the window and rises to meet the filter's at the last step. By $\mathrm{sd}(w_k)$, the steady gain is 0.00636, 0.01628, 0.06241 and 0.14813 at 0.01, 0.02, 0.05 and 0.10 K, and the time to 1.05 of the limit is 1320, 930, 420 and 210 s: the gain and the time are set by the rows chosen.

**The vendor's gain as rows.** The block's update has the form of equation (9.15) with a constant gain and no input. A constant gain is a steady limit, so the block can be asked which rows it is the limit of. With $a = 1$ and no load row, equation (9.14) gives $P^- = P + q$, and equation (9.15) gives $P = (1 - \upsilon)P^-$, so $q = \upsilon P^-$; with $P^- = \upsilon R/(1 - \upsilon)$ this is

$$q = \frac{\upsilon^2 R}{1 - \upsilon}. \tag{9.22}$$

At $\upsilon = 0.05$ this is a spread of 0.02565 K per step, and equation (9.21) on those rows returns a gain of 0.050000 and an updated spread of 0.1118 K. **A gain on a sheet is a choice of rows that the sheet does not state.** The vendor's block says the hall does not relax toward anything and the load does not exist.

### 9.3.5 The hall as a virtual reading through the load step

**Control engineer's question.** What did the hall do through the load step?

**General form.** Moments, equation (2.10), under the marginal (9.19) for each $k$ in a stretch of steps, with $j = k$ for what could be known at the moment and $j = N$ for what can be known afterward.

**Specialized form.** Equations (2.10) and (9.19) with $\mathbf{z}_k = T_k$ at $k = 81, \ldots, 120$. The sensor reads $T_k$ with a spread of 0.50 K, so the belief is a virtual reading of the same entry, carried by the load row and the neighbors' rows. The comparison must be like with like: the raw reading, the block and the filter use readings through $k$, and the stack uses all of them.

**Building blocks.** The filter and the stack of Section 9.2.3; the block of Section 9.1.2; the record's truth, as a check only.

**The plant's numbers.** Figure 9.4 draws steps 70 to 120. The horizontal axis is the step and the vertical axis the hall air. The gray dots are the readings, the solid black line the stated truth, the red line the vendor's block, and the blue and gold shaded bands the filter's and the stack's means with two of their standard deviations either side. Table 9.6 gives the RMS error against the truth over the stretch before the step and the stretch through it, with one column for each estimate.

Table 9.6: RMS error against the record's stated truth. The stack uses the readings after each step; the others do not.

| RMS [K] | Raw reading | Vendor's block | Filter | Stack |
|---|---|---|---|---|
| Steps 70–80, before the step | 0.4228 | 0.0282 | 0.0422 | 0.0115 |
| Steps 81–120, through it | 0.4391 | 0.7484 | 0.0915 | 0.0597 |

Before the step the block beats the filter on this record. Through it, at this gain, the block is worse than the raw reading, 0.956 K behind at step 100, while the filter's largest error is

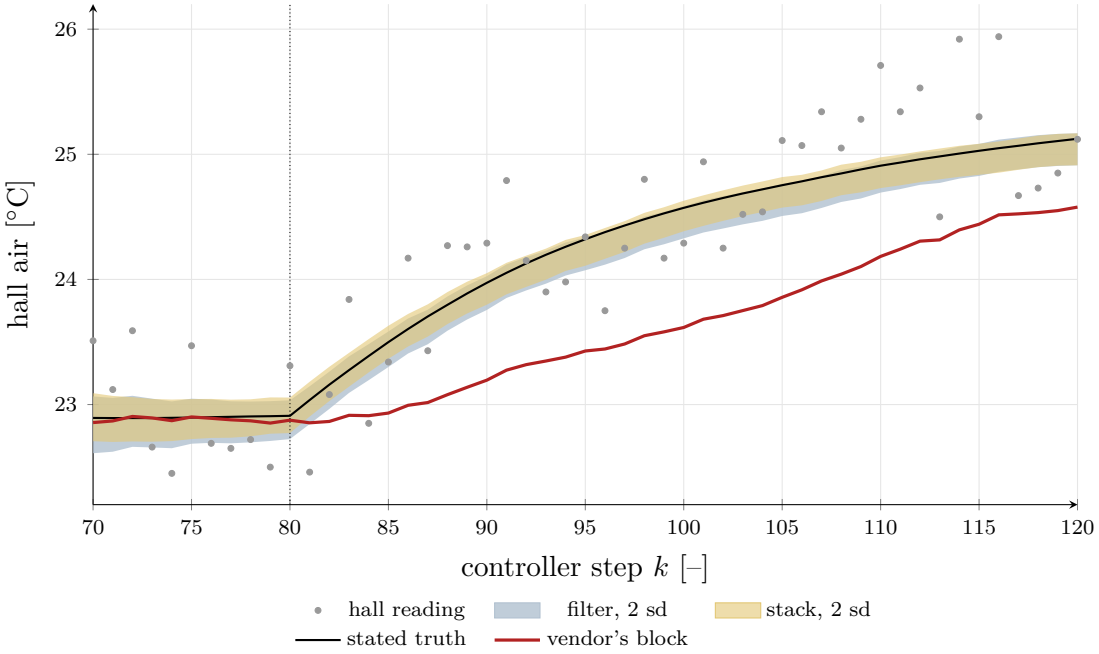



Figure 9.4: The hall air through the load step, known-$G$ rows, $T_c = r$. The dotted vertical line is step 80. The bands are the means of the filter and the stack with $\pm 2\sqrt{P_{k|k}}$ and $\pm 2\sqrt{P_{k|120}}$. The record is constructed from the stated truth.

0.130 K, at step 97. The block's 0.0801 K was never a band on the hall: it is the noise of its output on a hall that does not move. Its gain nearly equals the row's $1 - a = 0.0502$, a lag of 570 s against the hall's $C_h/G = 582$ s. Table 9.7 runs the block at four gains. Each row is a gain; the columns give its lag and its noise on a still hall, the RMS error on this record, and over 400 further records, seeds 30000 to 30399, the mean RMS error and the share of records on which the block beats the raw reading, whose RMS error there is 0.4958 K. "Worse than raw" holds only at this gain, and the best row there is still about twice the filter's 0.1222 K on the same records. The vendor's gain placed on the filter's row and input, with $a$ and the load kept and the gain held at 0.05, gives 0.0676 K on this record. **A gain trades lag for noise; the load row removes the lag without that trade.**

Table 9.7: The block by its gain, steps 81–120, one record and 400 further records.

| Gain [–] | Lag [s] | Noise [K] | RMS, this record [K] | RMS, 400 records [K] | Beats raw [%] |
|---|---|---|---|---|---|
| 0.05 | 570 | 0.0801 | 0.7484 | 0.7928 | 0.0 |
| 0.10 | 270 | 0.1147 | 0.4334 | 0.4794 | 57.2 |
| 0.20 | 120 | 0.1667 | 0.2528 | 0.2796 | 100.0 |
| 0.30 | 70 | 0.2100 | 0.2353 | 0.2468 | 100.0 |

Over 200 fresh records of the same plant, seeds 9000 to 9199 with Chapter 5's fit redone on each, the RMS error is 0.4916 K raw, 0.7957 K for the block and 0.1256 K for the filter, which beats the block on 100.0 % of them.

**The limit in $T_c$ and $G$.** Table 9.8 runs the same filter with the three pairs of $T_c$ and $G$, the last known to no fit. Each row is a pair; the columns give its values, the belief at step 120, whose truth is 25.124°C, the error in the belief's own standard deviations, and the RMS error through the step.

Table 9.8: The filter at step 120 by the pair in its input row. Err: estimate minus truth, in the row's own sd. In the stated truth, the 17.70°C is the steady value $r - b$; over this window the supply air has a mean of 17.76°C and ranges from 17.65 to 17.91°C.

| Pair | $T_c$ [°C] | $G$ [kW/K] | $T_{120}$ (sd) [°C] | Err [sd] | RMS 81–120 [K] |
|---|---|---|---|---|---|
| Chapter 5, $T_c = r$ | 18.0000 | 6.1827 | 25.040 (0.0639) | −1.32 | 0.0915 |
| Chapter 6, $T_c = r - \hat{b}$ | 17.9429 | 6.1110 | 25.055 (0.0641) | −1.07 | 0.0841 |
| Truth's pair, unknown to any fit | 17.7000 | 5.8000 | 25.137 (0.0653) | +0.21 | 0.0477 |

**The identity holds and the band is too narrow.** The pair $(T_c, G)$ is a false pin: Chapter 5's $G$ partly cancels the offset at 30 kW, and the cancellation fails at 45 kW. On this record its steady node stands −0.020 K from the truth's at 30 kW and −0.180 K at 45 kW. Figure 9.5 shows it over the fresh records. The horizontal axis is the error of $T_{120}$ in the belief's own standard deviations, the bars count records per bin of width 0.5, and the dotted lines mark ±2 standard deviations. The left panel pins $G$ and reads $T_c = r$, with the records whose normalized innovations exceed their 0.95 level drawn red. The right panel frees $g$, on the grid route of Section 9.3.6.

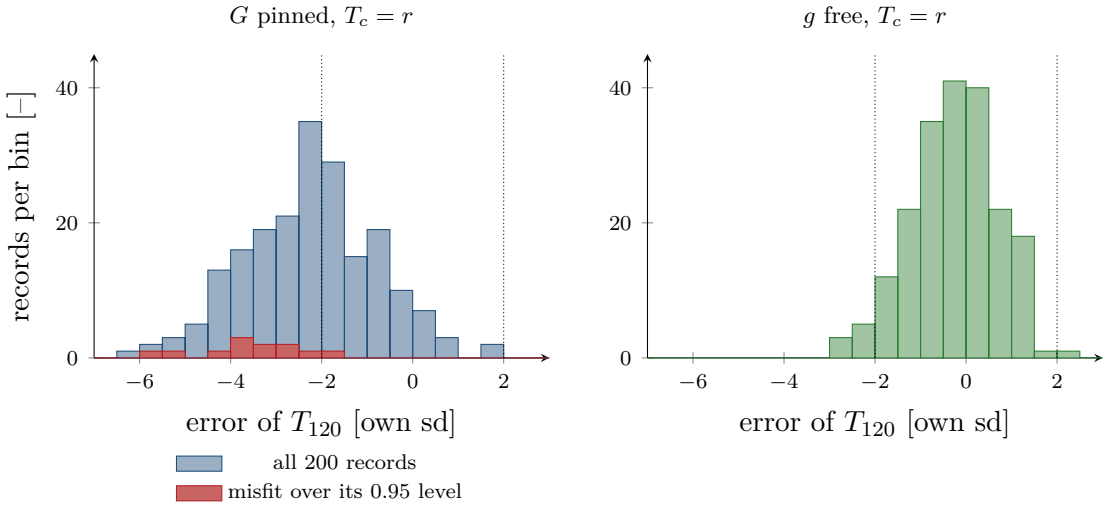



Figure 9.5: The error of $T_{120}$ in the belief's own sd over 200 fresh records, seeds 9000 to 9199. Left: inside 2 sd on 42.5 % (se 3.5), and the misfit over its level on 6.0 %. Right: inside 2 sd on 95.5 % (se 1.5). All percentages are this example's.

With Chapter 5's pair pinned the error has a mean of −2.25 of its own standard deviations, and only 42.5 % of the records (se 3.5) fall inside two. The misfit does not see this: $\sum_k \nu_k^2/S_k$ exceeds its 0.95 level on 6.0 % of the records, about what chance would give, and of those twelve records, eleven lie outside two standard deviations, so nearly all the records outside two pass the test. On this record it is 65.77 on 61 readings, tail probability 0.315. A weaker sign is there: over steps 81 to 120 the mean innovation is +0.1766 K, with a standard deviation of 0.0798 K for that mean, $z = +2.21$ here, and it is above +2 on 28.0 % of records (se 3.2). This is Chapter 5's finding on the coil, a misfit passed by a fit that is wrong, and here it falls on the belief. The right panel is what the question of Section 9.3.6 repairs.

**The pair and not the pin.** Table 9.9 keeps $G$ pinned in every row and changes the pair, over the 400 further records. Each row is a pair; the columns give the mean error of $T_{120}$ in

its own standard deviations, the share inside two with its standard error, and the RMS error through the step. Chapter 5's pair repeats the 42.5 % as 45.0 %. The truth's $G$ with $T_c = r$ does worse, the truth's $T_c$ with Chapter 5's $G$ far worse, and the truth's pair, still pinned, covers. The shortfall therefore comes from a wrong pair and not from pinning $G$. That $g$ is biased, not merely uncertain: over these records it has a mean of +0.0275 and a spread of 0.0182, its stated standard deviation is 0.0170, and the truth is −0.0339. A prior at its band shows the cost: with $g$ free on that prior, $T_{120}$ is inside two standard deviations on 90.0 % of the records (se 1.5) and $g$ on 3.0 % (se 0.9).

Table 9.9: Pinned pairs over 400 further records, seeds 30000 to 30399, sd($w_k$) = 0.02 K, $T_{120}$ against the truth in the row's own sd.

| $G$ | $T_c$ | Err mean [sd] | Inside 2 sd [%] (se) | RMS 81–120 [K] |
|---|---|---|---|---|
| Chapter 5's | $r$ | −2.11 | 45.0 (2.5) | 0.1222 |
| Truth's | $r$ | +2.47 | 19.5 (2.0) | 0.1368 |
| Truth's | $r - 0.30$ | −0.77 | 98.2 (0.7) | 0.0671 |
| Chapter 5's | $r - 0.30$ | −5.53 | 1.5 (0.6) | 0.2992 |

**The choice of** sd($w_k$)**.** Table 9.10 runs the recursion at seven values, with each row a value of sd($w_k$). The columns give the steady gain, the time to 1.05 of the limit, and, on this record at Chapter 5's pair, the spread and error of $T_{120}$, the RMS error through the step and the misfit.

Table 9.10: The stated choice sd($w_k$) = 0.02 K among others, $T_c = r$, $G$ pinned.

| sd($w_k$) [K] | Steady gain [–] | Time [s] | sd $T_{120}$ [K] | Err [sd] | RMS 81–120 [K] | $\sum \nu^2/S$ [–] |
|---|---|---|---|---|---|---|
| 0.002 | 0.00278 | 1620 | 0.0270 | −4.81 | 0.1150 | 67.72 |
| 0.005 | 0.00358 | 1530 | 0.0305 | −4.16 | 0.1134 | 67.59 |
| 0.010 | 0.00636 | 1320 | 0.0402 | −2.90 | 0.1081 | 67.16 |
| 0.020 | 0.01628 | 930 | 0.0639 | −1.32 | 0.0915 | 65.77 |
| 0.050 | 0.06241 | 420 | 0.1249 | +0.00 | 0.0688 | 61.43 |
| 0.100 | 0.14813 | 210 | 0.1924 | +0.07 | 0.1262 | 56.77 |
| 0.200 | 0.30203 | 90 | 0.2748 | −0.22 | 0.2061 | 50.06 |

The value sd($w_k$) = 0.05 K looks better here, and over the fresh records it gives 95.5 % (se 1.5) inside two standard deviations and an RMS error of 0.1071 K against 0.1256 K, at the price of a band of 0.1249 K. That value was seen only after the fresh records and the truth, so it is reported and not adopted: the chapter's rows stay at the 0.02 K chosen beforehand, and the comparison stands as a measure of that choice.

The 42.5 % is therefore as much the choice of sd($w_k$) as the pair. Table 9.11 sets the two repairs side by side over the 400 further records. Each row is a value of sd($w_k$). The first two columns of numbers keep Chapter 5's pair pinned; the rest free $g$ on the handbook's prior by the grid, and give the mean of $T_k$ under the weights on $g$ at step $k$ and under those at the last step. Coverage is what separates the repairs: pinned it runs from 18.8 to 96.2 %, and with $g$ free it stays between 96.2 and 99.0 %. The RMS error does not separate them in $g$'s favor at the moment of reading, 0.1241 K at 0.02 K against 0.1043 K pinned at 0.05 K. The 0.0713 K uses weights from readings after step $k$, so it is for the log and not for the loop's next move.

### 9.3.6 The recursion with $g$ free

**Control engineer's question.** With $g$ free, does the recursion still give the marginal?

Table 9.11: The pin and the free $g$ by sd($w_k$), $T_c = r$, 400 further records, seeds 30000 to 30399. Inside: $T_{120}$ within 2 of its own sd, se in parentheses. RMS over steps 81–120.

| | Pair pinned | | $g$ free | | |
|---|---|---|---|---|---|
| sd($w_k$) [K] | Inside [%] | RMS [K] | Inside [%] | RMS, step's weights [K] | RMS, last weights [K] |
| 0.005 | 18.8 (2.0) | 0.1353 | 96.2 (0.9) | 0.1234 | 0.0704 |
| 0.010 | 24.0 (2.1) | 0.1321 | 96.5 (0.9) | 0.1235 | 0.0704 |
| 0.020 | 45.0 (2.5) | 0.1222 | 98.0 (0.7) | 0.1241 | 0.0713 |
| 0.050 | 96.2 (0.9) | 0.1043 | 99.0 (0.5) | 0.1312 | 0.0875 |

**General form.** Moments, equation (2.10), of entries and derived quantities under the marginal (9.19), now on rows that are not linear.

**Specialized form.** Equations (2.10) and (9.19) with $\mathbf{z}_k = (T_k, g)^\mathsf{T}$, $k = j = 120$, on the rows (9.16). The derived quantities are the process gain to the load $K_Q = 1/(G_{\text{ref}}\mathrm{e}^g)$, with $G_{\text{ref}} = 6.0\,\text{kW/K}$, and the hall at a closed-loop scenario node, with $r$ and $Q^* = 45\,\text{kW}$ pinned and the valve free, $T^* = T_c + Q^*K_Q$. By the delta method (4.53), with the gradient $-K_Q$ for $K_Q$ and $-Q^*K_Q$ for $T^*$ in the column of $g$, their spreads are $K_Q\sqrt{\Sigma_{gg}}$ and $Q^*K_Q\sqrt{\Sigma_{gg}}$. Three routes reach the marginal, and draws check the third:

- The *iterated stacked fit*, the state route, is the posterior of Section 9.2.4.
- The *one-pass extended filter* runs the elimination of Section 9.2.2 on $(T_k, g)^\mathsf{T}$. The dynamics row is linearized at the current mean, with $\mathbf{F}_k = \left[\begin{smallmatrix} a & a'(\hat{T}-T_c)+\phi' y_{Q,k} \\ 0 & 1 \end{smallmatrix}\right]$. Equation (9.4) with the matrix inversion identity of Section 3.5.3, $g$ carried by a row whose tolerance goes to zero, gives the predict $\mathbf{\Sigma}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{\Sigma}_{k-1|k-1}\mathbf{F}_{k-1}^\mathsf{T} + \text{diag}(q, 0)$, and the rank-one form (4.74) with $\mathbf{h} = (1, 0)^\mathsf{T}$ gives the update with $\mathbf{\Upsilon}_k = \mathbf{\Sigma}_{k|k-1}\mathbf{h}/(\mathbf{h}^\mathsf{T}\mathbf{\Sigma}_{k|k-1}\mathbf{h} + R)$.
- The *grid* uses the fact that, given $g$, the rows are linear. The recursion then gives $\hat{T}_{k|k}(g)$, $P_{k|k}(g)$ and the likelihood $p(\mathbf{y} \mid g) = \prod_k \mathcal{N}(\nu_k;\ 0,\ S_k)$, and the moments follow by summing over $g$ on a grid, with the total variance rule (3.32) for $T_{120}$.

**Building blocks.** The fit, $\mathbf{J}$ and $\mathbf{\Sigma}$ of Section 9.2.4; the recursion of Section 9.2.2; the prior on $g$.

**The plant's numbers.** Table 9.12 gives the routes on the same window at $T_c = r$, one row each, with the draws and the truth as rows; the columns are the entry, the two derived quantities and the entry $T_{120}$.

Table 9.12: The three routes with $g$ free, $T_c = r$, sd in parentheses, with reweighted draws of $g$ (seed 9, $2 \times 10^5$ draws, effective size 132 188) and the truth.

| Route | $g$ [–] | $K_Q$ [K/kW] | $T_{120}$ [°C] | $T^*$ at 45 kW [°C] |
|---|---|---|---|---|
| Iterated stack, 6 iterates, band from $\mathbf{J}^\mathsf{T}\mathbf{J}$ | −0.01501 (0.01977) | 0.16919 (0.00334) | 25.2579 (0.1156) | 25.613 (0.1505) |
| One-pass extended filter | −0.01221 (0.01920) | 0.16871 (0.00324) | 25.2451 (0.1136) | 25.592 (0.1458) |
| Grid, $g$ summed | −0.01496 (0.01981) | 0.16921 (0.00335) | 25.2575 (0.1157) | 25.615 (0.1509) |
| Reweighted draws | −0.01499 (0.01984) | 0.16922 (0.00336) | 25.2576 (0.1158) | – |
| Truth | −0.03390 | 0.17241 | 25.1238 | 25.459 |

**The recursion no longer gives the marginal, by a small amount here.** What separates the routes is when each row is linearized. The one-pass filter linearizes each row once, at the mean of its moment, and never returns; the stack relinearizes every row at the final $g$.

The size, in the grid's standard deviations and in the hall's own units: the one-pass filter's $g$ lies $+0.139$ standard deviations away, $K_Q$ is $-0.294\,\%$ against the grid, the node at 45 kW is $-0.0224$ K, and the spread of $g$ has a ratio of 0.9690. The $K_Q$ gap is 0.149 of its own band, so on this hall it does not change what the gain is known to be. The iterated stack lies 0.002 standard deviations from the grid in $g$. Across the fresh records the largest one-pass $g$ gap is 0.526 standard deviations, with mean $+0.055$; the largest $K_Q$ gap is 0.001743 K/kW, which is 0.0784 K at 45 kW. The iterated stack stays within 0.004 standard deviations of the grid on every record.

The gap grows with what the one-pass filter must linearize across. With the prior's mean kept and its spread widened, the $g$ gap is $+0.104$ at a prior spread of 0.1, $+0.139$ at 0.3, $+0.417$ at 1.0 and $+1.842$ at 3.0, where $K_Q$ is $-3.6\,\%$, 0.16327 against 0.16939 K/kW. The handbook's prior is what keeps the gap small here, and that is this example's result; the extended filter is not dismissed by it.

Table 9.13 carries the belief through time and Figure 9.6 draws its first three columns. In the figure the horizontal axis is the step and the vertical axis $g$. Each route has its mean and a bar of one standard deviation, offset slightly along the axis so that the bars can be told apart, and the truth is dotted. At each $k$ the stack is solved to that step, so all three use the same readings.

Table 9.13: The belief with $g$ free through the window, $T_c = r$, sd in parentheses. "Stack to $k$" is the iterated stacked fit solved through step $k$.

| | $g$ [–] | | | $T_k$ [°C] | | |
|---|---|---|---|---|---|---|
| $k$ [–] | One-pass | Grid | Stack to $k$ | One-pass | Grid | Truth [°C] |
| 62 | 0.12098 (0.28744) | 0.11617 (0.27450) | 0.12426 (0.28821) | 23.156 (0.2568) | 23.145 (0.2569) | 22.889 |
| 70 | 0.10449 (0.13333) | 0.08300 (0.15050) | 0.10083 (0.14738) | 22.747 (0.2314) | 22.753 (0.2319) | 22.893 |
| 80 | 0.03683 (0.06586) | 0.02536 (0.07345) | 0.02955 (0.07303) | 22.872 (0.1752) | 22.886 (0.1820) | 22.910 |
| 85 | 0.05114 (0.05183) | 0.04524 (0.05579) | 0.04722 (0.05572) | 23.390 (0.1604) | 23.399 (0.1654) | 23.499 |
| 90 | 0.00436 (0.04227) | −0.00489 (0.04625) | −0.00371 (0.04592) | 23.976 (0.1505) | 23.997 (0.1574) | 23.973 |
| 100 | 0.01647 (0.03028) | 0.01277 (0.03177) | 0.01304 (0.03176) | 24.511 (0.1352) | 24.522 (0.1385) | 24.572 |
| 110 | −0.00968 (0.02346) | −0.01368 (0.02454) | −0.01363 (0.02443) | 25.000 (0.1228) | 25.016 (0.1265) | 24.909 |
| 120 | −0.01221 (0.01920) | −0.01496 (0.01981) | −0.01501 (0.01977) | 25.245 (0.1136) | 25.258 (0.1157) | 25.124 |

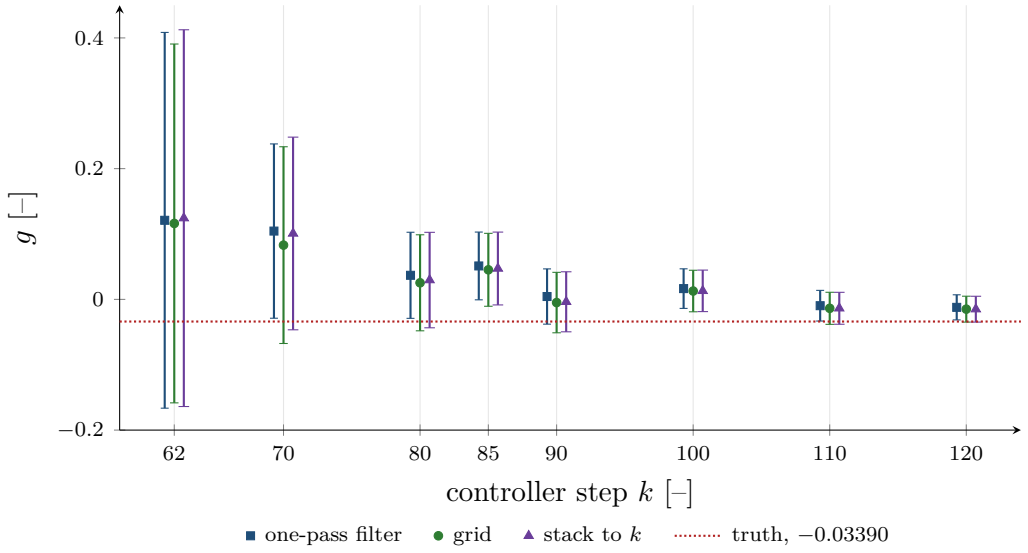



Figure 9.6: The belief on $g$ through the window by three routes, each with a bar of one sd, at the steps of Table 9.13. The routes are drawn 0.7 of a step apart at each $k$.

The figure shows the bars closing over the same stretch by all three routes. The largest differences are early: at step 70 the one-pass filter's mean is 0.10449 against the grid's 0.08300, with a narrower bar, and by step 120 the gap is the one quoted above. The gain on $T$ is not yet near a limit at step 120: 0.05166 with $g$ free against 0.01632 with it known, because the hall's readings are still teaching $g$, at $-0.007171$ per kelvin of innovation.

**Which $T_c$, with $g$ free.** Reading $T_c$ warmer than it is passes into $G$, the sign of Chapters 5 and 6, not re-derived here. Table 9.14 repeats the fit by the grid route with each input row. Each row is a choice; the columns give $g$, $K_Q$, $b$ when it is an entry, $T_{120}$, and the errors, estimate minus truth, in the row's own standard deviations.

Table 9.14: The input row with $g$ free, by the grid route, so that the rows compare like with like. The oracle row uses the truth's offset and no fit could. Beside the last row, the iterated stack gives $g = -0.01576$ (0.03810), $b = 0.0045$ (0.1957) and $T_{120} = 25.258$ (0.1170).

| Input row | $g$ (sd) [–] | $K_Q$ [K/kW] | $b$ (sd) [K] | $T_{120}$ (sd) [°C] | Err $g$, $b$, $T$ [sd] |
|---|---|---|---|---|---|
| $T_c = r$ | −0.01496 (0.01981) | 0.16921 | – | 25.258 (0.1157) | +0.96, –, +1.16 |
| $T_c = r - 0.0571$ | −0.02444 (0.01975) | 0.17082 | – | 25.263 (0.1159) | +0.48, –, +1.20 |
| $T_c = r - 0.30$, oracle | −0.06384 (0.01949) | 0.17769 | – | 25.285 (0.1164) | −1.54, –, +1.38 |
| $b$ free, prior 0.20 K | −0.01431 (0.03819) | 0.16919 | −0.0008 (0.1954) | 25.257 (0.1172) | +0.51, −1.54, +1.14 |

This record flatters the first row. Its +0.96 on $g$ is mild: over the fresh records $T_c = r$ has a mean error of +1.94 on $g$, with 54.5 % inside two standard deviations, while $b$ free gives 97.5 %. On $T_{120}$, both give 95.5 % (se 1.5), against the 42.5 % with Chapter 5's pair pinned. **Freeing $g$ repairs the belief on $T$; it does not repair $g$ without $b$.**

Set against Chapter 5, the window with the load step reads $G$ at 5.91 kW/K against 6.18, with $T_c = r$ in both. A rise of $T$ with $Q$ reads $1/G$ free of $T_c$, but only in part: $g$ still moves 0.04887 between $T_c = r$ and $r - 0.30$. Set against Chapter 6, there is a refinement. With $g$ flat and the offset's prior of 0.20 K kept, steps 40 to 79, one operating point, keep 100.00 % of $b$'s prior variance. Steps 60 to 120 keep 96.94 %, and steps 40 to 159 keep 85.57 %, as much as a stand-alone reading of 0.49 K. A load step opens that row a little. These shares are at $\mathrm{sd}(w_k) = 0.02$ K, this window: across 0.002 to 0.10 K they run from 96.59 to 99.08 % and from 82.22 to 97.31 %. Chapter 10 reads the same line with more rows.

## 9.4 The checks (step 5)

The answers above have already set the routes side by side. This section tests the recursion at the two ends of its rows, and tests the routes against each other and against the record's stated truth.

### 9.4.1 The limits of the dynamics row

Taking $\sigma_w, \sigma_Q \to 10^{-5}$ makes the dynamics row a physics row in its limit, with the chain carrying one unknown forward deterministically. The variance at step 120 then has the closed form $a^{2M}/(1/s_0^2 + \sum_{i=0}^{M} a^{2i}/R)$, with $M = N - K_0 = 60$ and $s_0 = 0.50$ K. The filter gives $4.61172 \times 10^{-5}\,\mathrm{K}^2$ and the closed form $4.61162 \times 10^{-5}\,\mathrm{K}^2$, a ratio of 1.000022, which differs from one by the small tolerance left in. At the other end, $\sigma_w \to 10^4$ cuts the row, and the filter returns the reading and its variance: 25.1200°C against 25.12°C, and 0.25000 against $R = 0.25000\,\mathrm{K}^2$.

### 9.4.2 The routes against each other

Table 9.15 collects the comparisons, with one row per check. The columns name what is compared, the values and the result. The algebra is what makes the first three hold; the rest are agreements to a stated size, and no agreement between the grid and the draws proves either of them.

Table 9.15: The checks of the chapter.

| Check | Values | Result |
|---|---|---|
| Stack against filter, every $k$ | largest gaps $2.9 \times 10^{-13}$ K, $1.1 \times 10^{-15}$ K$^2$; tolerance $10^{-9}$ | holds, by algebra |
| Stack's diagonal against backward pass | $6.7 \times 10^{-13}$ K, $1.1 \times 10^{-16}$ K$^2$ | holds, by algebra |
| Load kept against eliminated | $1.5 \times 10^{-13}$ K, $7.4 \times 10^{-17}$ K$^2$ | holds, by algebra |
| Readings moved by +1 K: gain and variance | largest change 0.0 | the band is known before the readings |
| $\sigma_w, \sigma_Q \to 10^{-5}$ against the closed form | ratio 1.000022 | agrees |
| $\sigma_w \to 10^4$ against the reading | 25.1200; 0.25000 K$^2$ | agrees |
| Grid, 1601 against 401 points | mean $g$ $2.0 \times 10^{-16}$, sd $g$ $2.6 \times 10^{-16}$, $T_{120}$ 0.0 | below $10^{-14}$ |
| Grid about the fit against a grid from −4.74 to +4.59 prior sd, 5601 points | mean $g$ $6.1 \times 10^{-16}$, sd $g$ $2.8 \times 10^{-15}$ | below $10^{-14}$ |
| Draws against the grid, $g$ | −0.01499 against −0.01496, se 0.00005 | within one se |
| Draws against the grid, $T_{120}$ | 25.2576 against 25.2575, se 0.0003 | within one se |
| Iterated stack against the grid, fresh records | $g$ gap mean −0.003, largest 0.004; sd ratio mean 0.999 | the band from $\mathbf{J}^\mathsf{T}\mathbf{J}$ holds here |

### 9.4.3 The answers against the stated truth

The record is constructed, so its truth can be set against the answers, as a check on the method on its own model. At step 120 the truth is 25.124°C. The belief with $G$ pinned stands −1.32 of its standard deviations from it, and the belief with $g$ free +1.16. The truth's $g$ is −0.03390, against −0.01496 on the grid at +0.96. The fresh records, seeds 9000 to 9199 with the stated truth unchanged, are the wider test, and Figure 9.5 reports it; they are the reason this one record is not read on its own.

## 9.5 The report (step 6)

The filter is no separate device: on linear rows with normal errors it is the last step's marginal of the stacked fit, by algebra. The surprise is what that identity does not protect. The band the filter reports is only as wide as its pins, and with Chapter 5's pair pinned the truth lay inside two standard deviations on 42.5 % (se 3.5) of the fresh records, unseen by the misfit test. With $g$ free the one-pass filter differs from the stack by 0.294 % of the hall's gain on this record, which is small here. These percentages are this example's.

### 9.5.1 The answers and their bands

Table 9.16 gives the answers in the order of Table 9.1. Each row is a question. The second column gives the state route's answer with the other routes beside it, the third the band, and the last the record's stated truth.

Table 9.16: The report on the hall's window. State route first; other routes named.

| Control engineer's question | Answer | Band | Stated truth |
|---|---|---|---|
| Is the filter a separate device? | No, on linear rows: stack and filter at step 120, 25.039713°C | gaps $1.3 \times 10^{-13}$ K; tolerance $10^{-9}$ | – |
| What is the hall at the last step? | $G$ pinned, $T_c = r$: stack 25.040°C. $g$ free: stack 25.258°C, grid 25.258°C. $b$ free: stack 25.258°C, grid 25.257°C | sd 0.0639 K; 0.1156, grid 0.1157 K; 0.1170, grid 0.1172 K | 25.124°C |
| What do later readings say about earlier steps? | At step 85: stack 23.498°C, filter 23.442°C. RMS over steps 81–120: 0.0597 against 0.0915 K (one record) | sd ratio 0.450 at step 60, 0.917 at 80, 0.971 at 115 | 23.499°C at step 85 |
| How fast does the band fall, and to what gain? | Within 1.05 of the limit in 930 s; steady gain 0.01628 at $\mathrm{sd}(w_k) = 0.02$ K | sd 0.3536 to 0.0638 K; the gain is set by the rows | – |
| What did the hall do through the load step? | Virtual reading: filter RMS 0.0915 K, block 0.7484 K, raw 0.4391 K. Fresh records: 0.1256, 0.7957, 0.4916 K | sd 0.0639 K at step 120 with Chapter 5's pair pinned, inside 2 sd on 42.5 % of fresh records | record, Fig. 9.4 |
| With $g$ free, does the recursion still give the marginal? | Not to algebra. $K_Q$: stack 0.16919, grid 0.16921, one-pass 0.16871 K/kW; $b$ free: stack 0.16931, grid 0.16919 K/kW | sd 0.00334, 0.00335, 0.00324; $b$ free: 0.00645, 0.00646 K/kW | 0.17241 K/kW |

### 9.5.2 The entry in the commissioning log

The entry gives the hall at step 120 with both of its bands: 25.040°C with a spread of 0.0639 K with $G$ pinned, and 25.258°C with 0.1156 K with $g$ free, and it uses the wider. It states the rows behind the block's gain: a gain of 0.05 means $a = 1$, no load row and 0.02565 K per step, so the block's 0.0801 K does not stand as a band on the hall, and through the step the block was 0.956 K behind on this record. It says that the pair $(T_c, G)$ is not to be pinned in the filter the set-point layer reads; the PI row reads $y_{c,k}$. Chapter 5's process gain to the load of 0.16174 K/kW is replaced, with the offset free, by 0.16931 K/kW with a spread of 0.00645 (grid 0.16919, 0.00646), and the offset itself stays unresolved. The narrower 0.00334 of $T_c = r$ is not installed: it holds $g$ inside two standard deviations on only 54.5 % of the fresh records. The verdict rests on one record's draw of its sensors' errors, and on the stated $\mathrm{sd}(w_k)$.

### 9.5.3 Common mistakes

Table 9.17 lists five mistakes this chapter's plant makes easy. Each row names the mistake, its form on this hall and the correction.

## 9.6 The next measurement (step 7)

The rows at the stack's linearization, with $g$ and $b$ both free, say what each candidate would remove before it is read. Nothing is refit.

### 9.6.1 The candidates

A candidate with Jacobian line $\mathbf{J}_a$ gives the preposterior (4.73), here in the stack's columns. A hall sensor adds $1/\sigma_a^2$ on each $T_k$ diagonal element; a better load meter adds it on each $Q_k$; a discharge sensor reads the air directly and acts on the column of $b$, a permanent one with its own offset of 0.2 K combined with its noise. The entry the set-point layer needs is $T_{120}$, so the candidates are ranked by that spread, from the present 0.1170 K, with 0.00645 K/kW on $K_Q$

Table 9.17: Five mistakes in reading a filter.

| Mistake | On this hall | Correction |
|---|---|---|
| Buying a gain as a property of a device | 0.05 on the sheet, with $a = 1$, no load row and 0.02565 K unstated | ask for the rows; a gain is their steady limit, equation (9.21) |
| Reading a block's noise as a band on the plant | 0.0801 K against an RMS error of 0.7484 K through the step | noise on a still hall is not a band on a moving one; carry the load row |
| Taking a filter's spread at its word | 0.0639 K with Chapter 5's pair pinned; inside 2 sd on 42.5 % of fresh records | free what is really uncertain; a recursion's agreement with the stack says nothing about its pins |
| Clearing the pins by the misfit test | tail probability 0.315 on this record; over the level on 6.0 % of records | the misfit tests readings against rows; test the band against records or free the entry |
| Choosing sd($w_k$) after seeing the truth | 0.05 K looks better once the 95.5 % is known | state the choice before the comparison; report the better value without adopting it |

and 0.1957 K on $b$. Table 9.18 gives them, a row each, with the spreads after each candidate and the share of the spread on $T_{120}$ removed.

Table 9.18: The candidates, $g$ and $b$ free, at the stack's linearization, ranked by the spread on $T_{120}$. Percentages are this example's.

| Candidate | sd $T_{120}$ [K] | Removed [%] | sd $K_Q$ [K/kW] | sd $b$ [K] |
|---|---|---|---|---|
| Reference hall sensor 0.1 K, every step (permanent) | 0.0431 | 63.2 | 0.00539 | 0.1821 |
| Second hall sensor 0.5 K, every step (permanent) | 0.0909 | 22.3 | 0.00607 | 0.1936 |
| Load meter 0.2 kW, every step | 0.1154 | 1.3 | 0.00642 | 0.1957 |
| Reference discharge sensor 0.05 K, one reading | 0.1157 | 1.1 | **0.00361** | **0.0484** |
| Permanent second discharge sensor, its offset 0.2 K | 0.1163 | 0.6 | 0.00517 | 0.1400 |

More record has a different last step and is listed apart: to step 159, with the same sensors, the spread on $T_{159}$ is 0.0928 K and on $K_Q$ it is 0.00520 K/kW.

### 9.6.2 The recommendation

**The ranking depends on the entry.** The temperature wants a hall sensor, and the reference grade takes the band from 0.1170 to 0.0431 K if it stays in place. The gain wants the discharge reference, brought once, which nearly halves its band and takes the offset's from 0.1957 to 0.0484 K, and does almost nothing for $T_{120}$. So the recommendation is by what the set-point layer will read: a better hall sensor for the virtual reading, and the discharge reference for the process gain. As the permanent instrument, the second hall sensor at the existing grade removes 22.3 % here, against 0.6 % for a permanent second discharge sensor.

### 9.6.3 What the analysis showed

The sheet offered a device with a gain and a band. The analysis found no device: the block is a recursion on rows the sheet does not state, and the recursion on the hall's own rows is the commissioning fit with its oldest entries removed, by algebra, so that the tab with the fit and the block with the gain are one object with two sets of rows. A gain trades lag for noise; the load row removes the lag without that trade. The harder finding of the window lies elsewhere.

The identity held at every step while the band was too narrow on most records, because the pair $(T_c, G)$ the rows held fixed partly cancels the offset at 30 kW and the cancellation fails at 45 kW, and the misfit test passed nearly all of them. On this hall the distance between the recursion and the stack is not what matters. What matters is a false pin. The engineer takes back a question to put to any filter: what does it hold fixed, and what happens to its band when that is freed.

### 9.6.4 Limits of the model

**The stated choices.** The value of $\mathrm{sd}(w_k)$ and the prior on $T_{60}$ are stated choices. White noise cannot represent a persistent error of $G$ and $T_c$. Chapter 5's fit overlaps this window, a double use of steps 60 to 79 that no rows here correct: with $G$ pinned through $G$ and the prior on $T_{60}$, and with $g$ free through that prior alone, since only the prior on $g$ refuses the fit.

**The supply air.** $T_c$ is held over each window, and its transient, which ranges from 17.65 to 17.91°C in the truth, is in no row.

**One node.** This is the hall node alone, with no coil or valve row. The input is read, and whether a move's effect on the record matters is the question of Chapter 24.

**The comparisons by RMS.** The fixed-gain and $\mathrm{sd}(w_k)$ comparisons by RMS error are this hall's. A fixed gain of 0.05 on the filter's row does better than the chosen filter over the fresh records, 0.1078 against 0.1256 K, because the rows hold a bias: at the truth's pair, over the further records, it gives 0.0710 against the filter's 0.0671 K. That is no argument for fixed gains.

**The constructed record.** This is one window, and its percentages are this example's. The record is constructed from a stated truth, and the vendor's block with it, so the agreement tests the method on its own model and not the model against a real hall. What one reading does to a belief is the question of Chapter 10.

## 9.7 What to retain

**Definition 9.1** (The Kalman filter as a time-stacked fit)**.** On linear rows with normal errors, the Kalman filter is the elimination of a time-stacked fit, oldest entry first. Eliminating an entry is the predict, equation (9.14), and adding the next reading's row is the update, equation (9.15), so that its moments at each step are the last entry's marginal of the stack through that step, by algebra. Its $(\hat{\mathbf{z}}_{k|k}, \boldsymbol{\Sigma}_{k|k})$ is the Fit line of the problem (7.1), here with a metered load for its input.

Three conclusions carry beyond the hall, as statements of form and not of size. A filter's gain is the steady limit of some rows, so asking for the gain is asking for the rows. The agreement of a recursion with its stack says nothing about the rows, so a band may be too narrow while every gap is at rounding, and a misfit test may pass it. And with rows that are not linear, the recursion and the stack differ by when they linearize: how much depends on what the one-pass filter must linearize across.

Table 9.19 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 9.19: The terms of the chapter and where each is defined. Units for the hall's window.

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| Belief on step $k$ from $j$ | $\hat{\mathbf{z}}_{k\|j}$, $\mathbf{\Sigma}_{k\|j}$, eq. (9.1) | mixed | the book's time notation | 9.1.1 |
| Step of the air node | eq. (9.2) | K | the row between neighbors | 9.1.2 |
| Eliminating an entry | eq. (9.4) | – | the Schur complement of the information | 9.1.4 |
| Row variance | $q$, eq. (9.6) | $\mathrm{K}^2$ | the load eliminated | 9.1.4 |
| Predict | eq. (9.14) | K, $\mathrm{K}^2$ | an entry eliminated | 9.2.2 |
| Update, gain, innovation | eq. (9.15) | K, –, $\mathrm{K}^2$ | a reading's row added | 9.2.2 |
| Iterated stacked fit | eq. (9.16) | mixed | all rows, relinearized at the fit | 9.2.4 |
| Smoother | eq. (9.20) | K, $\mathrm{K}^2$ | the stack's diagonal, by a backward pass | 9.3.3 |
| Steady gain | eq. (9.21) | – | the limit of the recursion | 9.3.4 |
| Rows of a stated gain | eq. (9.22) | $\mathrm{K}^2$ | what a sheet's gain implies | 9.3.4 |
| Virtual reading | $\hat{T}_{k\|k}$, with its band | °C | an entry from the rows, not the sensor | 9.3.5 |
| One-pass extended filter | $\mathbf{F}_k$, $\mathbf{\Upsilon}_k$ | mixed | each row linearized once | 9.3.6 |

## 9.8 Exercises

**Exercise 9.1.** *Formulation.* A tank's level is read every step and its inflow is metered, while its outflow is known. Write the stacked rows and the known-input row. (With $A$ the tank's area and $\Delta t$ the step, the dynamics row is $(L_{k+1} - L_k - (\Delta t/A)(F_k - O_k))/\sigma_w$, the meter row $(F_k - y_{F,k})/\sigma_F$, the reading row $(L_k - y_{L,k})/\sigma_L$, and there is a prior row on the first level. By equation (9.4) with $x = F_k$, the inflow leaves the known input $(\Delta t/A)(y_{F,k} - O_k)$ and the variance $q = \sigma_w^2 + (\Delta t/A)^2\sigma_F^2$, with $a = 1$.)

**Exercise 9.2.** *Derivation.* Derive the predict and the update from the Schur complement. (The block (9.11) and equation (9.4) give the information $1/q - (a^2/q^2)\,Pq/(q + a^2P) = 1/(a^2P + q)$ and the right side $(a\,m + c)/(a^2P + q)$, so the prediction has mean $a\,m + c$ and variance $a^2P + q$. Adding the reading's row, precisions add, $1/P_{k|k} = 1/P_{k|k-1} + 1/R$, and the correction form (3.29) gives $\upsilon_k = P_{k|k-1}/(P_{k|k-1} + R)$.)

**Exercise 9.3.** *Derivation.* Show why the vendor's gain gives $\sigma_w = 0.02565\,\mathrm{K}$. (With $a = 1$, $P^- = P + q$ and $P = (1 - \upsilon)P^-$, so $q = \upsilon P^-$, and $\upsilon = P^-/(P^- + R)$ gives $P^- = \upsilon R/(1 - \upsilon)$. Then $q = \upsilon^2 R/(1 - \upsilon)$, with $\upsilon = 0.05$ and $R = 0.25\,\mathrm{K}^2$, and $\sqrt{q} = 0.02565\,\mathrm{K}$ per step.)

**Exercise 9.4.** *Application.* Run one step of the recursion from step 60, by hand. (From $\hat{T}_{60|60} = 22.94587°\mathrm{C}$ and $P_{60|60} = 0.125000\,\mathrm{K}^2$, the predict gives $0.949783 \times 22.94587 + 1.14433 = 22.93792°\mathrm{C}$ and $0.113227\,\mathrm{K}^2$. The gain is $0.113227/(0.113227 + 0.25) = 0.311725$, the innovation $23.55 - 22.93792 = 0.61208\,\mathrm{K}$, and the update is $23.12872°\mathrm{C}$ with variance $0.077931\,\mathrm{K}^2$.)

**Exercise 9.5.** *Application.* Find the steady gain at $\mathrm{sd}(w_k) = 0.05\,\mathrm{K}$. (By equation (9.6), $q = 2.565972 \times 10^{-3}\,\mathrm{K}^2$, with $a = 0.949783$ and $R = 0.25\,\mathrm{K}^2$. Then $\beta = R(1 - a^2) - q = 2.191230 \times 10^{-2}\,\mathrm{K}^2$, and equation (9.21) gives $P^- = 0.016640\,\mathrm{K}^2$, so the gain is $0.016640/(0.016640 + 0.25) = 0.06241$.)

# Chapter 10

# Feedback as Conditioning

## 10.1 One new reading and the state the loop acts on

The commissioning engineer has the loop's trend open. The discharge reading scatters 0.10 K about its set point from step to step, and the valve follows it. The hall reading of step 39, the last before the fitted window and not used by the fit, stands at 22.89°C. Two requests lie on the commissioning sheet. One offers a filter on the discharge reading, with the note that the noise will come out of the valve. The other asks whether another day of hall readings will open the back-off of 0.2565 K at 45 kW.

The sheet treats these as one question, more data, and the engineer would conclude that both numbers fall as data arrive: the scatter the loop acts on and the margin it stands behind. Neither the trend nor the sheet says what a reading does to anything except the number it replaces. The question of this chapter is

*what one new reading does to the state the loop acts on, and to the loop.*

**Jargon: innovation, predictive variance, rank-one update, frozen row, refit, schedule of the back-off, floor, shared offset, scalar filter, valve travel, noise-to-air gain, crossover**

The *innovation* is a reading less what the belief predicted for it, and the *predictive variance* is the variance of that difference before the reading comes. A *rank-one update* moves the belief by one scalar row without refitting. A *frozen row* keeps its Jacobian line at the point where the belief was fitted; a *refit* solves all the rows again. The *schedule of the back-off* (the preposterior covariance over the count of readings) is the back-off against the number of further readings, known before any of them is taken. The *floor* is what the schedule tends to as the count grows. A *shared offset* is one error common to every reading of a sensor. The *scalar filter* is a one-entry fit of a reading on a random walk, with a gain $L$ per step. *Valve travel* is the sum of the valve's absolute moves per hour. The *noise-to-air gain* is the magnitude of the response of the true supply air to the discharge sensor's noise at one frequency, and the *crossover* is the frequency at which the magnitude of the open loop falls through one.

**Key idea**

One scalar reading is a rank-one move of the belief: the band narrows by an amount known in advance, and the mean moves by an amount that is not. So the back-off has a schedule and a floor, and the floor belongs to one load. A fixed PI row takes in none of this. It uses at most a point estimate and never the covariance. At the pinned gains of this hall a filter that estimates the reading better leaves the true supply air worse, except the filter that knows the move, which is in effect a lower gain. With the gains retuned a filter is a trade among travel, scatter and peak, and no route dominates.

### 10.1.1 The update as a move of the belief

Chapter 4 added a second reading to a posterior in two forms. Its precision form adds the row's line to the information, and its rank-one form, equation (4.74), moves the mean by a gain vector times the innovation and subtracts an outer product from the covariance. Chapter 7 placed this step in the problem (7.1) as the line with Fit. Chapter 9 derives it for a state moving in time, and its scalar filter is used here and not derived again. This chapter studies what the same step does on the closed-loop hall, where the state barely moves at one condition and the row of a controller stands between a reading and the plant.

What the loop changes is the destination of an estimate. In a record with no controller a filtered reading changes a report and nothing else. Here the estimate enters the PI row, the row moves the valve and the valve moves the air, so a change in the estimate is a change in the plant. This is coupling 2 of Section 7.1.3. Coupling 4 of that section has the covariance enter the move, and Chapter 6's back-off is its limit route. Whether the PI row exercises it is one of the findings below.

### 10.1.2 The closed-loop hall and its belief (step 0)

The plant is the closed-loop hall of Chapter 6, on the reduced state of equation (6.5), $\mathbf{z} = (Q,\ g,\ g_w,\ b)^{\mathsf{T}}$, with the dependent quantities of equation (6.6). A what-if is the closed-loop scenario node of equation (6.7): the set point $r$ and the load $Q^*$ are pinned, the valve is free and determined by the controller's row. Further readings at the same condition are the record's steps 0 to 39, which Chapter 5 and Chapter 6 did not use.

The loop study runs in time, step $k$ being 30 s, on the sampled loop of the same hall with the PI gains pinned at $K_p = 0.10$ per kelvin and $K_i = 0.002$ per kelvin second. The schedule of the record is kept: 30 kW, then 45 kW from step 80, then $r = 17°\mathrm{C}$ from step 160, where the valve goes to its stop and is pinned again.

### 10.1.3 The control engineer's questions (step 1)

Table 10.1 lists the six questions. The first column is the question in the engineer's words, the second the general form it takes, the third its mathematical form on this hall, the fourth its category, the fifth the computation that answers it and the last the subsection. Here $\iota$ is the innovation and $S$ the predictive variance.

### 10.1.4 The evidence as rows (step 2)

The rows of Chapter 6's fit stay. Each further reading adds one line, with these modeling choices.

**Further readings as independent rows.** A further hall reading has standard deviation 0.5 K and is taken as independent of the others given the state, 30 s apart. This is what makes

Table 10.1: The questions asked of one more reading on the closed-loop hall. $T^*$ is the node of equation (6.7); $L$ is the gain of the scalar filter.

| Control engineer's question | General form | Mathematical form | Category | Computation | Answered in |
|---|---|---|---|---|---|
| What does one reading do to the belief? | moments, (2.10) | $(\hat{\mathbf{z}}, \boldsymbol{\Sigma}) \to (\hat{\mathbf{z}} + \mathbf{K}\iota,\ \boldsymbol{\Sigma} - S\mathbf{K}\mathbf{K}^{\mathsf{T}})$ | estimation | posterior | 10.3.1 |
| What is known before the reading comes? | moments of a preposterior | $\boldsymbol{\Sigma}_n$ against $n$ | estimation | preposterior | 10.3.2 |
| What can more hall readings not remove? | the same, in the limit | $\lim_n \mathrm{sd}\, T^*$ | estimation | preposterior | 10.3.3 |
| What does one reading do to the window? | threshold at a scenario node, (2.12) | $r_{95}$ and the window after $y$ | prediction under a threshold | posterior | 10.3.4 |
| Should the PI row act on the reading or on a filtered estimate? | expected consequence, one fixed law, (2.23) | travel, $T_c$, margin against $L$ | design | expected consequence | 10.3.5 |
| What do a step response's rows add? | moments, (2.10) | $p(g_w, b \mid \text{stacked rows})$ | estimation | posterior | 10.3.6 |

$n$ readings equal to one reading of $0.5/\sqrt{n}$, and it fails when the load wanders over several steps; the wander is not studied here beyond Chapter 5's amendment (Section 10.6.4).

**A sensor's offset as one entry.** A permanent second discharge sensor of noise 0.10 K has an offset of its own with a tolerance of 0.20 K. Carrying the offset as an entry lets $n$ readings average the noise and leave the offset; writing them as independent rows averages both (Section 3.5.7). The chapter computes both and says what the second claims.

**A second load as a new entry.** Readings at 45 kW add a load entry and rest on the same path $g$ and the same offset $b$ at both loads. This is what lets two loads read the line of Section 10.3.3, and it fails if the path changes with the load.

**Which row acts on which entry.** With $*$ for an entry that can be nonzero, the candidate lines in the columns $Q, g, g_w, b$ and, for the rows at a second load, its load entry $Q_2$ are

$$\begin{bmatrix} * & * & 0 & * & 0 \\ * & 0 & * & * & 0 \\ * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} \text{hall reading at 30 kW} \\ \text{valve reading at 30 kW} \\ \text{load meter at 30 kW} \\ \text{reference discharge reading} \\ \text{the loop's own discharge reading} \\ \text{hall reading at the second load} \\ \text{load meter at the second load} \end{matrix} \tag{10.1}$$

Two facts are read here. The loop's own discharge reading is the zero row of Chapter 5: any number of them adds nothing. And the hall row at 30 kW holds $Q$ and $b$ in the same combination at any count, so what it cannot separate more readings of it cannot separate either; the hall row at the second load has a different column of $g$, and that is the line that changes the answer.

**The PI row and the filter.** The loop study uses a fixed law, with the gains above and an integrator that holds at the stop. The law's input is an estimate,

$$c_k = u_0 + K_p e_k + I_k, \qquad I_{k+1} = I_k + K_i\,\Delta t\, e_k, \qquad e_k = \hat{x}_k - \hat{b} - r, \tag{10.2}$$

where $\hat{b}$ is a constant the row subtracts (zero unless stated) and $\hat{x}_k$ is the scalar filter of the discharge reading $y_{c,k}$. On a random walk of the reading, Chapter 9's scalar form is

$$\hat{x}_k = \hat{x}_k^- + L\,(y_{c,k} - \hat{x}_k^-), \qquad \hat{x}_k^- = \hat{x}_{k-1} + K_\mathrm{m}\,\delta_k, \qquad \frac{L^2}{1-L} = \frac{\sigma_w^2}{\sigma_v^2}, \tag{10.3}$$

with $\sigma_w$ the walk's step and $\sigma_v = 0.10\,\mathrm{K}$ the reading's noise. $L = 1$ is the raw reading. Two families are studied. With $K_\mathrm{m} = 0$ the filter knows nothing of the valve. The filter that knows the move takes $K_\mathrm{m}$ as Chapter 5's process gain and $\delta_k$ as a copy of the actuator's last move, a prediction row of the kind Chapter 9 owns. $L$ is a stated choice here and not fitted, since the question is what each value does. The estimate is scored against what it estimates: the true reading mean $T_c + b$.

**The rows of a step response.** Section 10.3.6 stacks the rows of steps in time on the notation of Chapter 9. Its entries are the initial $T$, $T_c$ and $u$, then $g$, $g_w$ and $b$, a load level per segment and a load per step, with the commands known, each being a function of past readings.

## 10.2 The posterior of the state (step 3)

Every answer starts from one posterior, Chapter 6's, and moves it by further rows. The form of the move is set here; what it does is the subject of step 4.

### 10.2.1 The belief at the 30 kW window

The rows are those of equation (6.9), with objective $\mathcal{L} = \frac{1}{2}\mathbf{r}^\mathsf{T}\mathbf{r}$ and the normal equations (6.10), whose information matrix of normalized rows is $\mathbf{J}^\mathsf{T}\mathbf{J}$; with unnormalized rows it is $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$. The reduced state is cited and not derived again. The fit gives

$$\hat{\mathbf{z}} = \begin{bmatrix} 30.05809 \\ 0.01833 \\ -0.16764 \\ 0.05708 \end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 0.02496 & 0.00088 & 0.00076 & -0.00027 \\ 0.00088 & 0.00177 & -0.00183 & -0.00732 \\ 0.00076 & -0.00183 & 0.00235 & 0.00913 \\ -0.00027 & -0.00732 & 0.00913 & 0.03605 \end{bmatrix}, \tag{10.4}$$

with spreads $0.15798\,\mathrm{kW}$, 0.04210, 0.04846 and $0.18986\,\mathrm{K}$, and $\mathrm{corr}(g_w, b) = 0.99242$. At $r = 18°\mathrm{C}$ and $Q^* = 45\,\mathrm{kW}$ the gradient of the node is $\mathbf{g}_{T^*} = (0,\ -7.36377,\ 0,\ -1)^\mathsf{T}$. So $T^*$ has mean $25.3067°\mathrm{C}$, spread $\sqrt{\mathbf{g}_{T^*}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}_{T^*}} = 0.1559\,\mathrm{K}$ and back-off $1.6449 \times 0.1559 = 0.2565\,\mathrm{K}$, which puts $r_{95}$ at $19.9368°\mathrm{C}$.

### 10.2.2 One further row, in two forms

A further scalar row has value $\mathbf{h}^\mathsf{T}\mathbf{z}$ plus normal noise of variance $\sigma^2$. Information adds, and in the form of Chapter 4,

$$\boldsymbol{\Sigma}'^{-1} = \boldsymbol{\Sigma}^{-1} + \frac{\mathbf{h}\mathbf{h}^\mathsf{T}}{\sigma^2}, \qquad \hat{\mathbf{z}}' = \boldsymbol{\Sigma}'\Big(\boldsymbol{\Sigma}^{-1}\hat{\mathbf{z}} + \frac{\mathbf{h}\,y}{\sigma^2}\Big). \tag{10.5}$$

Inverting the first by the matrix identity gives the rank-one form (4.74),

$$\hat{\mathbf{z}}' = \hat{\mathbf{z}} + \mathbf{K}\iota, \qquad \boldsymbol{\Sigma}' = \boldsymbol{\Sigma} - S\,\mathbf{K}\mathbf{K}^\mathsf{T}, \qquad \mathbf{K} = \frac{\boldsymbol{\Sigma}\mathbf{h}}{S}, \qquad S = \mathbf{h}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{h} + \sigma^2, \qquad \iota = y - \mathbf{h}^\mathsf{T}\hat{\mathbf{z}}. \tag{10.6}$$

Here $\mathbf{K}$ is the gain vector, which Chapter 9 writes $\boldsymbol{\Upsilon}_k$, clear of the controller's gains. Neither $S$ nor $\mathbf{K}$ contains $y$.

The hall reading's line at $\hat{\mathbf{z}}$ is $\partial T/\partial \mathbf{z} = (0.16364,\ -4.91869,\ 0,\ -1)$ and the reference sensor's is $(0,\ 0,\ 0,\ -1)$. The hall row is not linear in the state, through $\mathrm{e}^{g}$, so the rank-one form holds for it only with the line frozen.

### 10.2.3 Three routes for a nonlinear row

Three routes are named and kept apart. *Route A* is the rank-one form (10.6) on the row frozen at $\hat{\mathbf{z}}$. *Route B* takes the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ as a prior and relinearizes the new row by iterating on it, so it fits the rows

$$\mathbf{r}_B(\mathbf{z}) = \begin{bmatrix} \boldsymbol{\Sigma}^{-1/2}(\mathbf{z} - \hat{\mathbf{z}}) \\ (h(\mathbf{z}) - y)/\sigma \end{bmatrix}. \tag{10.7}$$

*Route C* is a refit: every row of equation (6.9) and the new one, with the curvature term of equation (6.11) in its covariance. Route C is the reported value wherever a value depends on the reading. Route A gives the schedule, because it is the route on which the covariance does not depend on the value.

## 10.3 The answers (step 4)

Each answer reads the posterior (10.4) moved by the rows of step 2, in the order of Table 10.1.

### 10.3.1 What one reading does to the belief

**Control engineer's question.** What does one reading do to the belief?

**General form.** Moments, equation (2.10), of the state under the posterior with one row more, which for a normal posterior and a linear row is equation (10.5).

**Specialized form.** Equation (2.10) under the posterior (10.5), in the form (10.6), with $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ of equation (10.4) and $\mathbf{h}$ the reading's line. A load-meter reading and a reference discharge reading are linear in the state, so routes A and C can differ only through the earlier rows.

**Building blocks.** $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$; $\mathbf{h}$, $\sigma$ and $S$ of the reading; a singular value decomposition of the drop $\boldsymbol{\Sigma} - \boldsymbol{\Sigma}'$, to see its rank.

**The plant's numbers.** A reference discharge reading of 0.05 K at 17.70°C has innovation $-0.24292$ K and predictive variance $0.038548\ \mathrm{K}^2$, with gain $(0.007,\ 0.18992,\ -0.23686,\ -0.93515)^{\mathsf{T}}$. It moves $\hat{b}$ to 0.28425 K and takes its spread from 0.18986 to 0.04835 K. A hall reading of 0.5 K has $S = 0.256236\ \mathrm{K}^2$ and gain $(0.00005,\ -0.00488,\ -0.00008,\ -0.00032)^{\mathsf{T}}$, with $\mathbf{g}_{T^*}^{\mathsf{T}}\mathbf{K} = +0.0363$. At the record's 22.89°C, innovation $+0.0284$ K, it leaves $\hat{b} = 0.05707$ K, with $T^*$ at 25.3077°C. So the hall reading of the opening scene barely touches the offset, and a reference takes most of its band away.

Figure 10.1 draws this in the plane of $(g_w, b)$. The horizontal axis is the offset and the vertical axis the coil's log-derate. The solid contour is the belief now, at one standard deviation, and the lighter one at two; the dashed contour, on top of the solid one, is the belief after one hall reading at 22.89°C; the purple contour is the belief after a reference reading at 17.70°C; the cross is the stated truth.

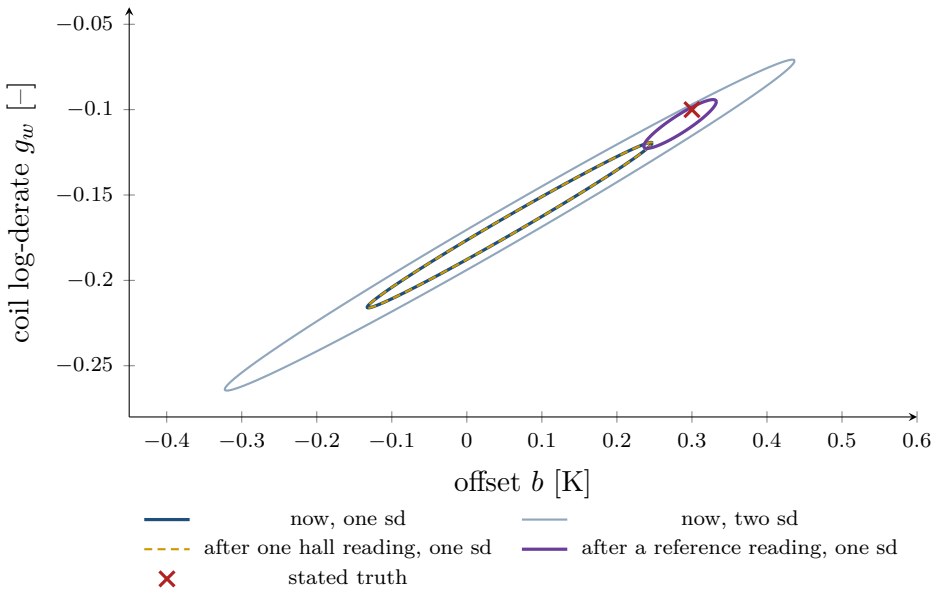



Figure 10.1: The belief in the plane of $(g_w, b)$ before and after one reading, each the refit's Laplace normal. The contours are the points $(\hat{b} + k\sqrt{\Sigma_{bb}}\cos t,\ \hat{g}_w + k\sqrt{\Sigma_{g_w g_w}}(\rho\cos t + \sqrt{1-\rho^2}\sin t))$: spreads 0.04846, 0.18986 and correlation 0.99242 now and after the hall reading, to the digits printed; 0.01429, 0.04833 and 0.90897 after the reference.

A reference reads the one column that the ridge of Chapter 6 leaves open, and the ellipse shrinks along the ridge. A hall reading has a zero in the column of $g_w$ and a gain of only $-0.00008$ and $-0.00032$ on these two entries, and the ellipse does not visibly change.

**The rank of the move and the two forms.** On the two linear rows the rank-one form and the precision form agree to at most $4.54 \times 10^{-12}$ of an entry's standard deviation in the mean, against a stated tolerance of $10^{-9}$. The drop in the covariance has a largest singular value of $3.72651 \times 10^{-2}$ on the reference row and a second of $2.5 \times 10^{-18}$, so it has rank one to that precision. Four frozen rows taken in all 24 orders reach the batch result within $2.28 \times 10^{-14}$ of an entry's standard deviation in the mean, so order does not matter. The covariance after a load reading of 30.5 kW and after one of 25 kW have a largest entry gap of 0.0: the value of a reading is absent from it.

**Relinearization.** Table 10.2 tests route A against route C. Each row is a reading; the columns give A less C on the mean of $T^*$, on its spread and on $r_{95}$.

Table 10.2: Route A, the rank-one form on the frozen row, less route C, the refit, at $r = 18°\text{C}$, $Q^* = 45\,\text{kW}$.

| Reading | $T^*$ [K] | sd $T^*$ [K] | $r_{95}$ [K] |
|---|---|---|---|
| Hall, $-2$ predictive sd | +0.0001 | −0.0013 | +0.0020 |
| Hall, $+2$ predictive sd | +0.0001 | +0.0013 | −0.0022 |
| Hall, the record's $k = 39$ | 0.0000 | 0.0000 | −0.0001 |
| Valve, the record's $k = 39$ | 0.0000 | 0.0000 | 0.0000 |
| Reference 0.05 K, at 17.70°C | +0.0079 | +0.0058 | −0.0175 |

The largest gap is on a row that is linear in the state. Its own gradient has no error, but the earlier rows are not linear and the reading moves the belief far along them: $g_w$ moves +4.15

of route C's spread. That entry is not the mechanism, since $T^*$ has no gradient on $g_w$, and A less C on it is only $-0.12$ of the spread. With the hall row linearized at $\hat{\mathbf{z}}$ and the rest refitted, $T^* = 25.4269°\mathrm{C}$, about route A's 25.4272°C; with the valve row linearized, 25.4192°C, about route C's 25.4193°C. So the gap is the hall row's curvature: A less C on $g$ is $-0.00104$, which gives $+0.0077$ of the $+0.0079\,\mathrm{K}$. So "linear rows" has to mean all of them. Route A's spread here uses the node's gradient at the moved mean, and that moved gradient is most of its $+0.0058\,\mathrm{K}$; with the frozen gradient the reference gives $0.1266\,\mathrm{K}$ against route C's $0.1272\,\mathrm{K}$, a gap of $-0.0006\,\mathrm{K}$. Route B is named and not recommended. On one valve reading it widens the belief, sd $T^*$ 0.1598 against $0.1559\,\mathrm{K}$, since a Newton curvature on one row with a residual has no other rows to rebalance it, and route C on the same reading stays at $0.1559\,\mathrm{K}$.

### 10.3.2 What is known before the reading comes

**Control engineer's question.** What is known before the reading comes?

**General form.** The preposterior, equation (4.73), of a candidate row, here $n$ of them, and the moments of a derived quantity under it,

$$\boldsymbol{\Sigma}_n = \left(\boldsymbol{\Sigma}^{-1} + n\,\mathbf{h}\mathbf{h}^{\mathsf{T}}/\sigma^2\right)^{-1}, \qquad \mathrm{sd}_n\, f = \sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_n\mathbf{g}}. \tag{10.8}$$

**Specialized form.** Equation (10.8) with $\mathbf{h}$ the hall line at 30 kW, $\sigma = 0.5\,\mathrm{K}$ and $\mathbf{g} = \mathbf{g}_{T^*}$. The back-off is $1.6449\,\mathrm{sd}_n\, T^*$, by equation (6.3). On the frozen row this is a function of $n$ alone: the claim that a back-off has a schedule known in advance is algebra for a linear row and has to be measured for this one.

**Building blocks.** $\boldsymbol{\Sigma}$ and the two gradients of Section 10.2.1; the refit of route C, to test the frozen row at values $\pm 2$ predictive standard deviations from the predicted mean.

**The plant's numbers.** Table 10.3 gives the schedule. Each row is a count; the columns give the time at 30 s a reading, the frozen spread, the back-off, $r_{95}$ if the mean stays where it is and the refit's spread at the two extreme means. Figure 10.2 draws it. In panel (a) the horizontal axis is the count on a logarithmic scale, the left vertical axis the spread and the right vertical axis the back-off; the solid line is the schedule, the dashed line the floor of Section 10.3.3 and the dots the refit on the record's own readings. In panel (b) the same axis carries the record's mean of $T^*$ with its spread as bars; the dashed line is the mean now and the dotted line the stated truth.

Table 10.3: The schedule of the back-off under further hall readings at 30 kW, 0.5 K each, at $r = 18°\mathrm{C}$, $Q^* = 45\,\mathrm{kW}$.

| $n$ [–] | Hours [h] | sd $T^*$ [K] | Back-off [K] | $r_{95}$, mean held [°C] | Refit at $\pm 2$ predictive sd, sd $T^*$ [K] |
|---|---|---|---|---|---|
| 0 | 0.00 | 0.1559 | 0.2565 | 19.9368 | – |
| 1 | 0.01 | 0.1549 | 0.2547 | 19.9386 | 0.1548, 0.1549 |
| 10 | 0.08 | 0.1468 | 0.2415 | 19.9518 | 0.1466, 0.1470 |
| 40 | 0.33 | 0.1319 | 0.2170 | 19.9763 | 0.1316, 0.1322 |
| 160 | 1.33 | 0.1151 | 0.1893 | 20.0040 | 0.1147, 0.1155 |
| 640 | 5.33 | 0.1062 | 0.1748 | 20.0186 | 0.1057, 0.1068 |
| 2560 | 21.33 | 0.1034 | 0.1700 | 20.0233 | 0.1028, 0.1039 |

The hall row is not linear in the state, so the covariance does depend on the value. The dependence is at most $0.0009\,\mathrm{K}$ of back-off out to $n = 10240$, between the frozen schedule and

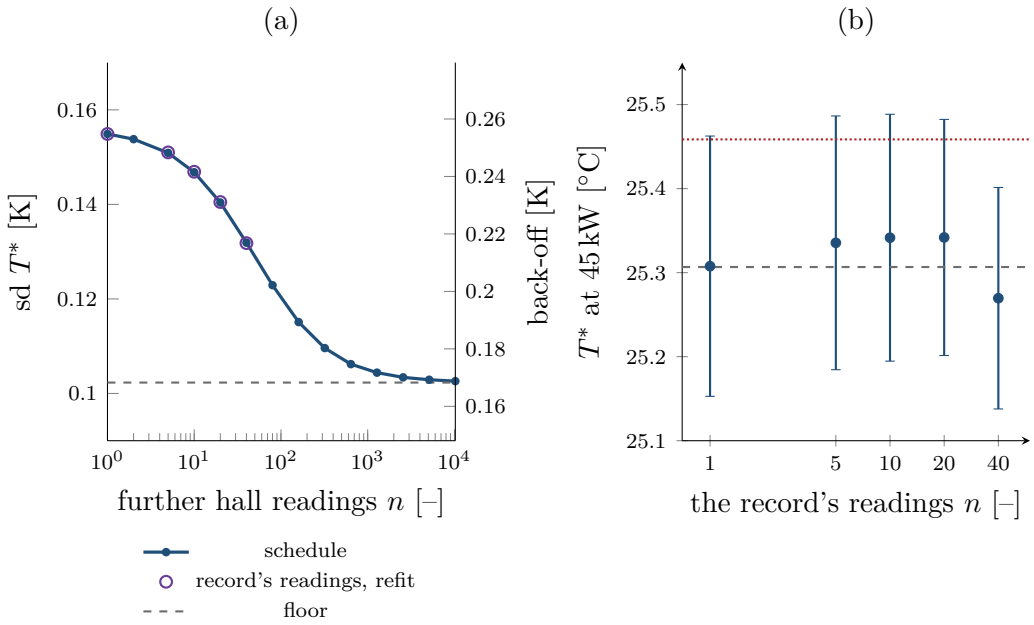



Figure 10.2: (a) The spread of $T^*$ and the back-off, 1.6449 times it, against the count: the frozen schedule of equation (10.8), the refit on the record's own readings and the floor 0.1023 K. (b) The record's mean of $T^*$ with its spread after the last $n$ readings of steps 0 to 39, the mean now, 25.3067°C (dashed), and the stated truth, 25.4586°C (dotted).

the refit at the extreme means. That is the tolerance on the claim. Panel (a) shows the record's own readings on the schedule: a spread of 0.1318 K at $n = 40$ against the schedule's 0.1319 K. Panel (b) shows what the schedule does not give. The mean walked: 25.3077, 25.3355, 25.3416, 25.3419 and 25.2696°C at $n = 1$, 5, 10, 20 and 40. The band was known. The mean was not, and it ended further from the truth, which stands +1.43 standard deviations from it.

That drift is not this record's alone. Over 300 fresh records of the record's own generator, seeds 424200 to 424499, forty further hall readings raise the mean distance of the truth from 0.793 to 0.921 stated standard deviations (standard error 0.049), and the share of records whose truth lies below the stated 95 % bound on $T^*$ falls from 0.840 to 0.820 (standard error 0.022). A second 300, seeds 24200 to 24499, gives 0.837 to 0.780 on that share. The prior on the offset has mean zero against a true 0.30 K, hall readings cannot correct it, and the band narrows all the same: the schedule is the stated band's, and it is known in advance whether or not it covers.

### 10.3.3 What more hall readings cannot remove

**Control engineer's question.** What can more hall readings not remove?

**General form.** The limit of equation (10.8) as $n$ grows, which is the covariance with the row's value known,

$$\boldsymbol{\Sigma}_\infty = \boldsymbol{\Sigma} - \frac{\boldsymbol{\Sigma}\mathbf{h}\mathbf{h}^\mathsf{T}\boldsymbol{\Sigma}}{\mathbf{h}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{h}}. \tag{10.9}$$

**Specialized form.** Equation (10.9) with $\mathbf{h}$ the hall line at 30 kW. The identity (6.15) says what remains: with $T$ at 30 kW known, $T^* = \lambda T - (\lambda - 1)r + (\lambda - 1)b$ and $\lambda = Q^*/Q$, so the spread left over belongs to $b$ and to $\lambda$ through the load. In the gradient on $(Q, b)$, the two entries are $-0.2450$ K/kW and 0.4971, with $\lambda = 1.4971$.

**Building blocks.** $\boldsymbol{\Sigma}$, $\mathbf{h}$ and $\mathbf{g}_{T^*}$ as before; the budget (4.54) on $(Q, b)$ under $\boldsymbol{\Sigma}_\infty$; the refit at $n = 10^5$ as a second route.

**The plant's numbers.** The floor is $\mathrm{sd}\,T^* = 0.1023\,\mathrm{K}$ and a back-off of 0.1683 K. Its variance of $0.01047\,\mathrm{K}^2$ splits into 0.00891 from $b$, 0.00150 from $Q$ and 0.00007 of cross term. The refit at $n = 10^5$ gives 0.1023 K. The spread of $b$ stays at 0.1899 K (refit 0.1898 K), set by the prior. Over the priors 0.05, 0.10, 0.20, 0.40 and 0.80 K, the floor on $\mathrm{sd}\,b$ is 0.0498, 0.0987, 0.1898, 0.3295 and 0.4517 K, and the back-off floor is 0.0752, 0.1027, 0.1683, 0.2782 and 0.3772 K. This is the statement of Chapter 3 that a fixed offset sets a floor more windows cannot remove, on this plant.

**When more readings stop paying.** The answer depends on the rule, so two are stated, and each is a choice. Under the rule that a further reading should remove at least 0.001 K of back-off, more readings stop paying from $n = 17$, 8.5 min, with the back-off at 0.2337 K. Under a rule of closeness to the floor, the spread is within 10, 5 and 1 % of it from $n = 213$, 478 and 2598, which are 1.77, 3.98 and 21.65 h. The whole removable amount is 0.0882 K of back-off, and twenty minutes gives 0.0395 K of it. The trend's own three rows at every step, hall, valve and load meter, lower the floor to 0.0944 K and leave $\mathrm{sd}\,b$ at 0.1898 K.

**The shared offset.** Table 10.4 sets two ways of writing a permanent second discharge sensor beside each other, following Section 3.5.7. Each row is a count; the first pair of columns carries the offset as an entry, the second pair writes the same readings as independent rows of the combined standard deviation.

Table 10.4: A permanent second discharge sensor, noise 0.10 K, offset 0.20 K, at $r = 18°\mathrm{C}$, $Q^* = 45\,\mathrm{kW}$.

| | Offset carried as an entry | | Written as independent rows | |
|---|---|---|---|---|
| $n$ [–] | sd $b$ [K] | Back-off [K] | sd $b$ [K] | Back-off [K] |
| 1 | 0.1447 | 0.2361 | 0.1447 | 0.2361 |
| 10 | 0.1385 | 0.2336 | 0.0663 | 0.2116 |
| 40 | 0.1379 | 0.2334 | 0.0348 | 0.2065 |
| $10^7$ | 0.1377 | 0.2333 | 0.0001 | 0.2046 |

The second form overstates at every $n > 1$ and promises a zero the sensor cannot reach. The loop's own discharge reading is a zero row at any count, as equation (10.1) shows, with the spreads left at 0.1899 and 0.1559 K. A hall offset $d$ is a stated what-if, since the record's hall sensor has none. For a prior on $d$ of 0, 0.05, 0.10, 0.20 and 0.50 K, the spread of $T^*$ as $n$ grows is 0.1023, 0.1266, 0.1805, 0.3130 and 0.7139 K. At 0.10 K the back-off floor is 0.2970 K, above what is held today without any further reading. A hall offset of 0.10 K costs more than the discharge offset does: taking the discharge sensor's prior from 0.05 up to its tolerance of 0.20 K moves the same floor only from 0.0752 to 0.1683 K.

**The floor is a floor at one load.** From equation (6.6),

$$T - r = \frac{Q}{G_{\mathrm{ref}} \mathrm{e}^{g}} - b, \tag{10.10}$$

a line in $Q$ whose slope is the path and whose intercept is the offset. One load gives one point on that line, with any offset met by some slope. Two loads give the line. Hall and load-meter readings at 45 kW, on the frozen rows of equation (10.1), with the path taken as the same at

both loads, give sd $b$ of 0.1607, 0.1543 and 0.1514 K at $n = 40$, 160 and $10^7$, and sd $T^*$ at 45 kW falls to 0.0734, 0.0402 and 0.0002 K, since the node is now read directly. Reading at both loads, $n$ steps at each, sd $b$ is 0.1498, 0.0688, 0.0187 and 0.0006 K at $n = 40$, 640, 10240 and $10^7$.

So the statement that no number of hall readings removes the offset's share is true at one load and false across two. Three conditions come with this: **the path must be the same at both loads, the hall sensor must be true, and the heat must be metered.** Each has a size. A path 1 % better at 45 kW shifts the intercept 0.1536 K; with a prior of 1 % on that difference, sd $b$ has a floor of 0.1168 K, and at 0.2 % of 0.0293 K. A meter offset or unmetered heat of 1 kW shifts it 0.1724 K. If the hall sensor has an offset $d$, the line's intercept reads only $b - d$, which is not fatal: a prior of 0.10 K on $d$ leaves sd $b$ and sd $T^*$ both at 0.0886 K. The plant's path is constant by construction, so the record cannot test any of this. And the sd $b$ of 0.0006 K takes 9.5 years at each load. Chapter 3's statement that the evidence is an independent measurement and not more windows of the same sensor stands for $b$ and the discharge sensor; the independent measurement here is the hall sensor, through the physics.

Chapter 9 read the same sensors, hall and load meter with $g$ flat, in time: steps 40 to 159 keep 85.57 % of $b$'s prior variance and steps 60 to 120 keep 96.94 %. Static rows of $40 + 80$ steps, read the same way, keep 65.8 %. That contradicts neither Chapters 3 and 6 nor that chapter's result, but the information beyond the prior differs about threefold, 13.000 against 4.216 per K$^2$, a ratio of 3.08. Static rows assume a settled hall. The hall's time constant is 20.7 steps, and after the load step the supply air's last excursion past 0.02 K is at 52 steps and the hall's at 89. So the two-load numbers here are those of settled rows, and the steps of the move itself do not earn them.

Figure 10.3 draws the line. The horizontal axis is the load and the vertical axis $T - r$. The blue line and band are the belief now, the purple line the belief with the record's steps 140 to 159, the dotted line the stated truth, and the circles the two window means. Panel (b) enlarges the left end, where the intercept is read as $-b$.

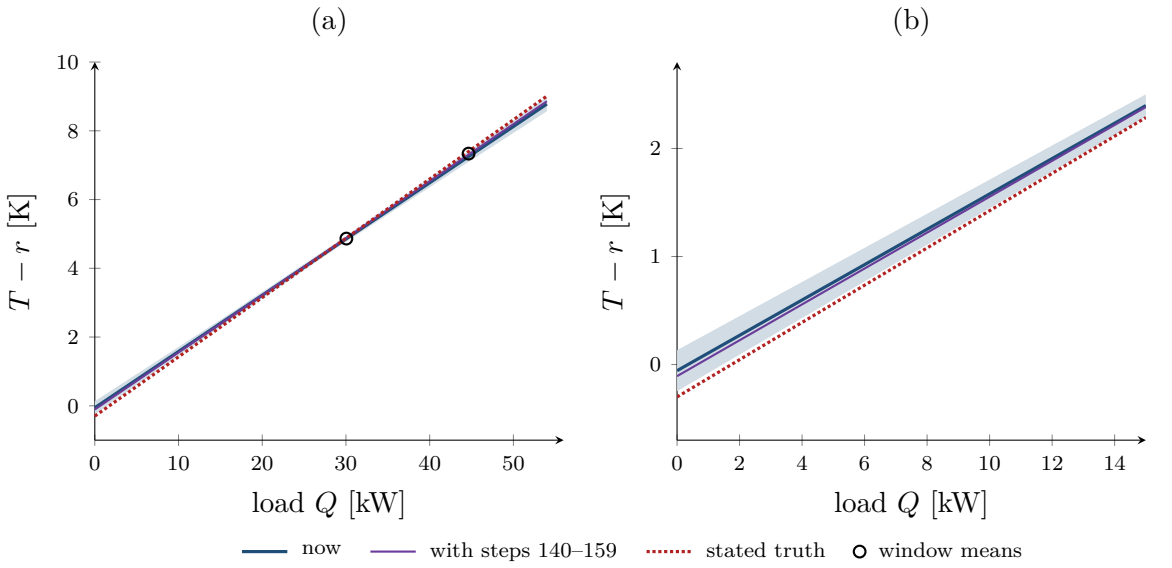



Figure 10.3: The line (10.10). The band is the belief now at one standard deviation, 0.1899, 0.1038, 0.0829, 0.1559 and 0.2122 K at 0, 15, 30, 45 and 54 kW, joined between those loads. Means and intercepts: now −0.0571 K, with the record's second load −0.1075 K, stated truth −0.3000 K. The load $Q = 0$ is an extrapolation of the line, read for its intercept and not a condition of the hall.

The band pinches near 30 kW, where the belief has its one point, and opens toward both ends. The record's own second load, steps 140 to 159, twenty readings with the hall still settling

(truth 25.330°C against 25.342°C steady at that load), gives $\hat{b} = 0.1075\,\mathrm{K}$ with spread 0.1658 K, $T^*$ of 25.3652°C with spread 0.0947 K and $r_{95} = 19.9790$°C. The truth stands +0.99 standard deviations away on $T^*$ and +1.16 on $b$.

### 10.3.4 What one reading does to the window

**Control engineer's question.** What does one reading do to the window?

**General form.** The threshold form (2.12) at a scenario node, inverted for its input, equation (6.17), under the posterior with the reading added.

**Specialized form.** Equation (2.12) through its inversion (6.17), with $f = T^*$ of equation (6.7) and the belief moved by equation (10.6). Then equation (4.69) holds on this plant, with $s$ and $s'$ the spreads of $T^*$ before and after and $c = \mathbf{g}_{T^*}^{\mathsf{T}}\mathbf{K}$:

$$\Delta r_{95}(\iota) = 1.6449\,(s - s') - c\,\iota. \tag{10.11}$$

The first term is the sure release. The second has mean zero under the predictive density of $\iota$ and variance $s^2 - s'^2$, by the covariance drop in equation (10.6), so the share of readings that lower $r_{95}$ is

$$\Pr(\Delta r_{95} < 0) = \Phi\left(-1.6449\sqrt{\frac{s - s'}{s + s'}}\right). \tag{10.12}$$

This rests on the frozen row, and the refit tests it.

**Building blocks.** $s$ and $\mathbf{g}_{T^*}$ of Section 10.2.1; $\mathbf{K}$ and $S$ of each reading; route C at five values of the reading; a 9-point Gauss–Hermite rule over the predictive density.

**The plant's numbers.** The window at 45 kW now runs from 17.438 to 19.937°C. Table 10.5 gives three readings. Each row is a reading; the columns give the spread after it, the sure release, the share of equation (10.12) and $r_{95}$ by route C at three values of the reading. Figure 10.4 draws those five-value curves in the manner of panel (b) of Figure 4.6. In its upper part the horizontal axis is the innovation in predictive standard deviations and the vertical axis $r_{95}$ after the reading; the dashed line is the level before and the dotted line the largest set point the stated truth allows. The lower part is the predictive density on the same axis.

Table 10.5: One reading and the set point at 45 kW, from $r_{95} = 19.9368$°C and sd $T^* = 0.1559$ K.

| Reading | sd $T^*$ after [K] | Sure release [K] | Share lowering $r_{95}$ [–] | $r_{95}$ at $-2, 0, +2$ predictive sd [°C] |
|---|---|---|---|---|
| One hall, 0.5 K | 0.1549 | 0.0018 | 0.4613 | 19.9754, 19.9386, 19.9018 |
| 40 hall readings | 0.1319 | 0.0395 | 0.3173 | 20.1431, 19.9763, 19.8095 |
| Reference 0.05 K | 0.1266 | 0.0482 | 0.2982 | 19.8014, 19.9850, 20.1683 |

On average the result of Chapter 4 holds. The Gauss–Hermite rule on the refit returns a mean $T^*$ of 25.3067°C, unchanged, and an $r_{95}$ gain of +0.0018, +0.0395 and +0.0481 K against the sure releases. At the center of the density each curve stands above the dashed line by that release.

Reading by reading it does not hold. Each curve crosses the dashed line, at +0.10, +0.48 and −0.53 predictive standard deviations, and the reference's slope has the opposite sign, because a higher reference reading lowers $\hat{b}$ and raises $\hat{g}$ along the ridge, and at 45 kW the better path outweighs the warmer air, $c = -0.4634$. A reference reading at the truth's 17.70°C narrows the

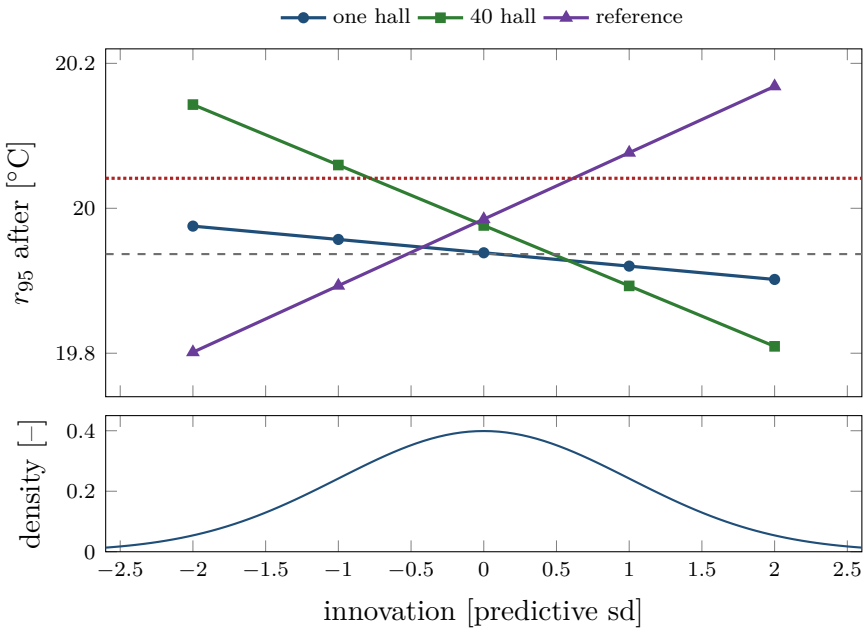



Figure 10.4: The largest set point at 45 kW after one reading, by route C, at the reading's mean and ±1, ±2 predictive standard deviations. Dashed: 19.9368°C before. Dotted: the stated truth's 20.0414°C. Below: the predictive density $\mathcal{N}(0, 1)$ of the standardized innovation.

band and moves $r_{95}$ from 19.9368 to 19.8714°C, while the true largest set point is 20.0414°C. Chapter 16 owns what that means for a set point. The spread after the refit stayed below its value before in every case tabulated, which is observed here and is a theorem only on linear rows. The window's lower edge moves only for the reference, over 17.362 to 17.473°C; a hall reading leaves it at 17.438°C. The record's own reading at $k = 39$, the last before the fitted window, moves $T^*$ to 25.3077°C and $r_{95}$ by +0.0008 K.

### 10.3.5 The PI row on the reading or on a filtered estimate

**Control engineer's question.** Should the PI row act on the reading or on a filtered estimate?

**General form.** The expected consequence (2.23) of an action, $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$, where the action is a choice of input to one fixed law and $L$ here lists several consequences side by side, with no weights chosen among them.

**Specialized form.** Equation (2.23) with $a$ the route, that is the value of $L$ in equation (10.3) and of $\hat{b}$ in equation (10.2), with the gains pinned. The consequences are the valve travel, the within-window standard deviation of the true $T_c$, the error of the estimate against $T_c + b$, the rise of $T_c$ after the load step, and the gain and phase margins. The expectation is over future noise and load at the stated truth, taken by fresh records, which buys a study of the loop itself and does not test model error. The records are 1000 in number, seeds 10000 to 10999, with common noise across routes, so each change is a paired difference. A second route takes the second moments of the linear map of the loop about the operating point.

The first thing to read is the row itself. Equation (10.2) has the estimate $\hat{x}_k$ and no other input from the fit. **The PI row never uses the covariance.** This holds by construction, since the row was written without it. The filter's gain is a constant, and no route here has the belief's $\mathbf{\Sigma}$ in the row; $L$ itself is a frozen covariance ratio of the walk model, equation (10.3), set once and not updated by any reading. So a fixed law does not run on the posterior in any sense

beyond its estimate: a sentence one is tempted to write, that the loop runs on the posterior and not on a point, is not supported by this computation. It uses at most a point estimate. The row does not exercise coupling 4 of Section 7.1.3; recomputing $r$ from the back-off would, by the limit route (Chapter 16); the loop study pins $r$.

**Building blocks.** The rows (10.2) and (10.3); the fresh records; the linear map, for the margins and for the noise-to-air gain. The linear map uses the truth's $(g, g_w, b)$, so the margins are the truth's and not the belief's; Chapter 8 carries the band of them.

**The plant's numbers.** Table 10.6 gives the routes at the steady window, steps 40 to 79 at 30 kW, and through the load step of step 80. Each row is a route. The columns give the travel with its standard error, the mean within-window standard deviation of the true $T_c$, the RMS error of the estimate, the peak rise of the mean trajectory of $T_c$ after the step, the mean square of $T_c$ about the pre-step mean over steps 80 to 119, and the two margins at 30 kW.

Table 10.6: The loop on the raw reading and on filtered estimates. $K_p = 0.10$, $K_i = 0.002$ above the rule; below it the stated $K_p$ with $K_i/K_p$ fixed, except the Kalman row at the pinned gains. 1000 fresh records, seeds 10000 to 10999, common noise. The standard errors of the other columns are at most 0.0007 K on the scatter, 0.0025 K on the peak and 0.00084 $K^2$ on the mean square. These are this hall's values at the stated gains.

| Route | Travel [/h] (se) | sd true $T_c$ [K] | Estimate error [K] | Load-step peak [K] | Mean square [$K^2$] | Gain margin [–] | Phase margin [deg] |
|---|---|---|---|---|---|---|---|
| Raw, $L = 1$ | 0.4112 (0.0018) | 0.0323 | 0.0998 | 0.1863 | 0.01400 | 14.39 | 60.7 |
| $L = 0.5$ | 0.2886 (0.0018) | 0.0342 | 0.0602 | 0.2129 | 0.01538 | 6.74 | 49.8 |
| $L = 0.3$ | 0.2398 (0.0020) | 0.0360 | 0.0493 | 0.2430 | 0.01783 | 4.65 | 39.2 |
| $L = 0.1$ | 0.1717 (0.0022) | 0.0415 | 0.0471 | 0.3386 | 0.03935 | 2.68 | 17.3 |
| $L = 0.05$ | 0.1272 (0.0019) | 0.0391 | 0.0489 | 0.4207 | 0.07279 | 1.81 | 7.0 |
| $L = 0.3$, knows the move | 0.1617 (0.0009) | 0.0220 | 0.0380 | 0.2869 | 0.04283 | 11.44 | 65.4 |
| $L = 0.1$, knows the move | 0.0622 (0.0004) | 0.0135 | 0.0224 | 0.4655 | 0.14634 | 20.64 | 67.8 |
| Raw, less $\hat{b} = 0.0571$ K | 0.4118 (0.0018) | 0.0328 | 0.0998 | 0.1854 | 0.01373 | – | – |
| Raw, less $\hat{b} = 0.2841$ K | 0.4142 (0.0018) | 0.0348 | 0.0998 | 0.1828 | 0.01276 | – | – |
| $L = 0.3$, $K_p = 0.0243$, raw's phase margin | 0.0558 (0.0004) | 0.0147 | 0.0435 | 0.4659 | 0.13670 | 19.12 | 60.7 |
| Raw, $K_p = 0.0243$ | 0.0979 (0.0004) | 0.0141 | 0.0998 | 0.4223 | 0.11819 | 59.12 | 68.9 |
| Raw, $K_p = 0.0641$, the peak of $L = 0.3$ | 0.2597 (0.0011) | 0.0245 | 0.0998 | 0.2430 | 0.02847 | 22.46 | 65.8 |
| Raw, $K_p = 0.0508$, the crossover of $L = 0.3$ knowing the move | 0.2049 (0.0009) | 0.0213 | 0.0998 | 0.2793 | 0.04137 | 28.35 | 67.3 |
| Kalman filter of the true plant, load walk 0.30 kW a step | 0.1044 (0.0009) | 0.0215 | 0.0288 | 0.3512 | 0.05780 | – | – |

**The result. At the pinned gains the filter estimates better and the loop runs worse, on everything except the valve.** At $L = 0.3$ and 30 kW the estimate's error halves, from 0.0998 to 0.0493 K, and valve travel falls 41.7 %, from 0.4112 to 0.2398 per hour. The true $T_c$ scatter rises 11.3 %, from 0.0323 to 0.0360 K, the paired change being +0.0037 K with a standard error of 0.0002 K. The load-step peak rises by a factor of 1.305, from 0.1863 to 0.2430 K, a paired change of +0.0568 K with a standard error of 0.0006 K. The margins fall from 14.39 and 60.7° to 4.65 and 39.2°. Over steps 80 to 119 the estimate's own error is 0.0588 K at $L = 0.3$

against 0.1012 K raw, and it turns: at $L = 0.1$ it is 0.1296 K, worse than the raw reading, since a filter that slow lags the move.

Figure 10.5 draws the load step. The horizontal axis is the number of steps after the load step at step 80 and the vertical axis the mean of the true $T_c$ over the 1000 records, less each record's mean over steps 40 to 79. Each curve is a route, and zero is the pre-step level.

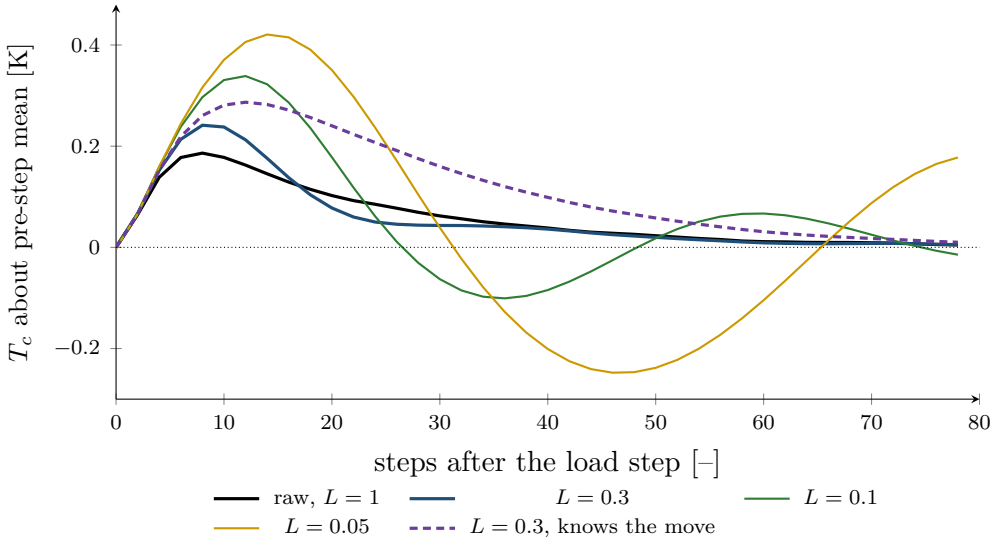



Figure 10.5: The mean trajectory of the true $T_c$ after the load from 30 to 45 kW, 1000 fresh records, seeds 10000 to 10999, $K_p = 0.10$, $K_i = 0.002$, plotted every two steps of 30 s.

The raw loop peaks first and lowest and settles without a swing. The filter that knows nothing of the valve raises the peak, and at $L = 0.1$ and $L = 0.05$ it rings: at $L = 0.05$ the mean swings to −0.2477 K and is still outside 0.02 K at the last step of the window. The dashed curve shows what the filter that knows the move pays: it rises higher and returns slowly, without ringing, as a loop of about half the gain does.

The comparison at 45 kW says the same thing with smaller sizes. Table 10.7 sets the two loads side by side on like measures. Each row is a route; for each load, the columns give the travel and scatter by the moment route over a 40-step window, the error of the estimate by records and the two margins. At 30 kW the error is over steps 40 to 79 and at 45 kW over steps 140 to 159, a stretch after the step.

Table 10.7: The two loads of comparison. Travel and sd $T_c$ by the second moments of the linear map over a 40-step window; estimate error from the 1000 fresh records. †: the mean trajectory has not settled by step 159, which inflates the error.

| | At 30 kW | | | | | At 45 kW | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Route | Travel [/h] | sd $T_c$ [K] | Error [K] | GM [–] | PM [deg] | Travel [/h] | sd $T_c$ [K] | Error [K] | GM [–] | PM [deg] |
| Raw, $L = 1$ | 0.413 | 0.0339 | 0.0998 | 14.39 | 60.7 | 0.405 | 0.0284 | 0.1003 | 21.87 | 66.9 |
| $L = 0.5$ | 0.291 | 0.0363 | 0.0602 | 6.74 | 49.8 | 0.280 | 0.0297 | 0.0602 | 10.68 | 57.8 |
| $L = 0.3$ | 0.243 | 0.0386 | 0.0493 | 4.65 | 39.2 | 0.229 | 0.0310 | 0.0484 | 7.23 | 48.1 |
| $L = 0.1$ | 0.172 | 0.0456 | 0.0471 | 2.68 | 17.3 | 0.154 | 0.0342 | 0.0477 | 4.18 | 25.2 |
| $L = 0.05$ | 0.127 | 0.0442 | 0.0489 | 1.81 | 7.0 | 0.113 | 0.0316 | 0.1372† | 3.04 | 13.8 |
| $L = 0.3$, knows the move | 0.163 | 0.0236 | 0.0380 | 11.44 | 65.4 | 0.159 | 0.0192 | 0.0390 | 17.46 | 70.4 |

At $L = 0.3$ and 45 kW the estimate's error falls from 0.1003 to 0.0484 K and the travel from 0.405 to 0.229 per hour, while the air's scatter rises from 0.0284 to 0.0310 K and the margins

fall from 21.87 and 66.9° to 7.23 and 48.1°. The records of steps 140 to 159 agree in direction and with less to show: a window of 20 steps gives a scatter of 0.0230 against 0.0237 K, a paired change of +0.0007 K with a standard error of 0.0002 K. At $L = 0.2$ that short-window change is +0.0001 K with a standard error of 0.0003 K, not resolved; the moment route gives 0.0322 K over its 40 steps. The load step itself runs from one load to the other, so its peak is not repeated in this table.

**Why.** The reading's noise is what moves the air: of the within-window spread of $T_c$ on the raw loop, the moment route gives 0.0337 K from that noise alone and 0.0043 K from the load alone. The noise reaches the air only below the loop's bandwidth, the crossover being 0.00718 rad/s at 30 kW, and down there the filter removes little. The valve's travel lies above it. Through $K_p$ the noise moves the valve at every frequency, and the filter's cut at high frequency is what takes the travel down. What the filter adds at the crossover is lag, and lag raises the noise-to-air peak. That peak is 1.076, 1.193, 1.495, 3.364 and 8.249 at $L = 1$, 0.5, 0.3, 0.1 and 0.05 at 30 kW, and 1.046, 1.084, 1.231, 2.304 and 4.195 at 45 kW. The filter's mean delay at $L = 0.3$ is 70 s, and it takes the delay margin at 30 kW from 147 to 106 s.

Figure 10.6 draws this. Each panel has the filter's gain $L$ on a logarithmic horizontal axis, the raw reading at the right end. The panels show the travel, the scatter of the true $T_c$ and the phase margin at 30 kW. The solid line is the filter with no knowledge of the move, and the squares are the filter that knows it. The dotted lines are the raw values.

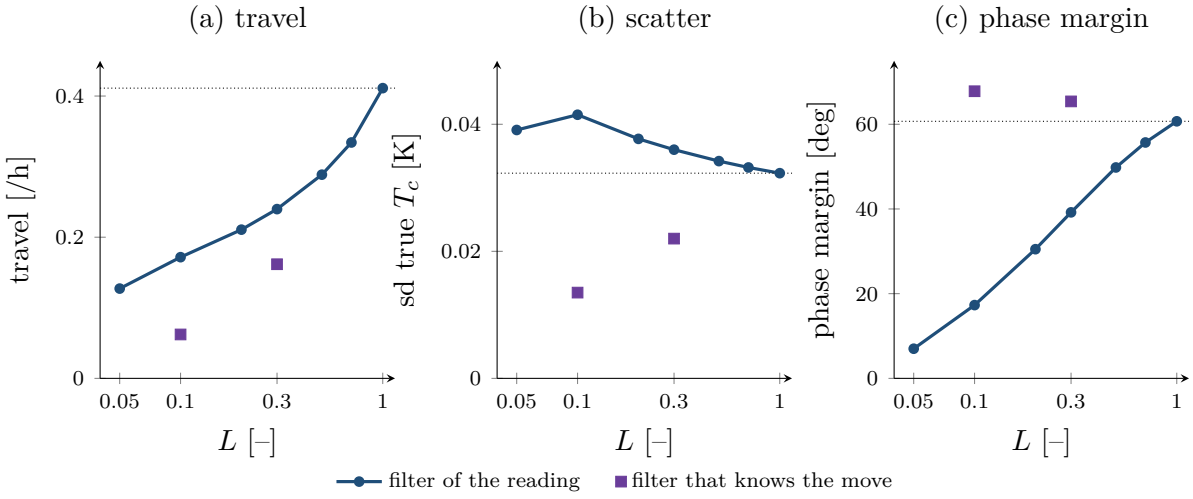



Figure 10.6: The routes against the filter's gain at 30 kW, $K_p = 0.10$, $K_i = 0.002$: valve travel and the mean within-window sd of the true $T_c$ over steps 40 to 79 of 1000 fresh records, seeds 10000 to 10999, and the phase margin of the linear map at the truth. Dotted: the raw reading.

Panel (a) is what the sheet promises, and it holds. Panels (b) and (c) are what the sheet leaves out. The dip of panel (b) at $L = 0.05$ comes from the record's quiet start: the slow oscillation has not yet built up by steps 40 to 79. The same 40-step window at stationarity, 200 records (seeds 20000 to 20199) at steps 400 to 439, shows no dip: 0.0445 K at $L = 0.1$ (standard error 0.0015) and 0.0562 K at $L = 0.05$ (standard error 0.0023). The stationary spread is 0.0372 K raw and 0.0418 K at $L = 0.3$, 0.0373 and 0.0418 K by the Lyapunov equation, so the ranking holds.

**The trap, and coupling 2.** Suppose the gain is taken by fitting a random walk to the quiet loop. The step-to-step standard deviation of the true air on the raw loop is 0.0118 K, and equation (10.3) gives $L = 0.1108$. At the nearest computed gain, $L = 0.10$, the phase margin is

17.3° and the load-step mean square is 2.81 times the raw value. The loop's quiet is the loop's doing, so a random walk fitted inside the loop calls for a filter that undoes the quiet.

**Knowing the move.** A prediction row with Chapter 5's process gain changes the pinned numbers on the scatter. At $L = 0.3$ it gives lower travel, lower scatter, 0.0220 against 0.0323 K, a better phase margin, 65.4°, and 3.06 times the load step's mean square. This is in effect a detune: the row halves the effective gain. The raw reading at the same crossover, 0.00396 rad/s, needs only $K_p = 0.0508$, and gives a scatter of 0.0213 K, a peak of 0.2793 K and a mean square of 0.04137 $\mathrm{K}^2$ against the row's 0.0220, 0.2869 and 0.04283. Only the travel is better, 0.1617 against 0.2049 per hour. Nor is the row a correct model. It applies $K_\mathrm{m} = -7.10$ in one step, where the plant's gains are $-2.16$ in one step and $-6.61$ steady at 30 kW, and $-4.40$ steady at 45 kW. A true model does not remove the cost: a steady Kalman filter of the plant with a load entry, the load a random walk of 0.30 kW a step, gives a scatter of 0.0215 K and a mean square of 0.05780 $\mathrm{K}^2$. Chapter 9 owns those rows, and they are shown here once. Its two gain margins differ by definition: 9.77 by the controller's gain, from a critical $K_p$ of 0.9769, where the internal model does not move, and 11.44 by the plant's.

**The offset.** Every filter leaves the mean of $T_c - r$ at $-0.300$ K, and the RMS of $T_c - r$ where it was, 0.3023 K raw and 0.3030 K at $L = 0.3$. A filter of one sensor estimates the reading, not the air. Subtracting a reference reading's estimate, $\hat{b} = 0.2841$ K from the constructed reading at the truth's 17.70°C, takes the mean to $-0.0159$ K and the RMS to 0.0422 K, and the hall air rises 0.282 K with it, 25.705 against 25.422°C at step 159.

**What would change the answer.** The verdict above stands only at the pinned gains, this hall and this filter. Three things could change it. The gains, which Chapter 14 re-tunes. On the grid computed, $K_p$ from 0.10 to 1.10 with $K_i/K_p$ fixed and $L$ from 1 to 0.1, 200 records (seeds 10000 to 10199), no cell improves the $T_c$ scatter by filtering, and the filter destabilizes first: $K_p = 0.80$ is stable raw, radius 0.9523, and unstable at $L = 0.5$, radius 1.0223. On that grid the filter does not always cut the travel either, 1.947 at $K_p = 0.40$, $L = 0.3$ against 1.928 per hour raw. But that grid only raises $K_p$, and a filter would be retuned the other way, to win back its margin. The lower rows of Table 10.6 go there. At $L = 0.3$ and the raw loop's phase margin of 60.7°, $K_p = 0.0243$, the travel is 0.0558 per hour, the scatter 0.0147 K, the peak 0.4659 K and the mean square 0.13670 $\mathrm{K}^2$, against the pinned raw loop's 0.4112, 0.0323, 0.1863 and 0.01400: the scatter reverses, $-54.5$ %, and the peak is $\times 2.50$. The raw loop at that $K_p$, at nearly the same crossover (0.00211 against 0.00209 rad/s), gives 0.0979, 0.0141 K and 0.4223 K, at a phase margin of 68.9°. The raw loop at the pinned filter's peak, $K_p = 0.0641$, gives a scatter of 0.0245 against 0.0360 K and a travel of 0.2597 against 0.2398 per hour, with a larger mean square, 0.02847 against 0.01783 $\mathrm{K}^2$. This is a trade among travel, scatter and peak, and no route dominates. What stands at every gain is that the PI row has no covariance in it. The filter: one that knows the move lowers the scatter as a lower gain does, as above. The law: a law derived with the belief, in Part IV, chooses its gains and its estimator together, and that is a different row from this one with a filter added.

The set-point step to the stop is the same on all routes, with the true $T_c$ at step 239 from 17.120 to 17.129°C: there the valve is pinned and neither the reading nor its filter acts.

### 10.3.6 What a step response's rows add

**Control engineer's question.** What do a step response's rows add?

**General form.** Moments, equation (2.10), of $g$, $g_w$ and $b$ under the stacked rows of a window, by the Laplace covariance of the stack.

**Specialized form.** Equation (2.10) with the state the stack of Section 10.1.4, the rows being the plant's exact step with its command known, on the notation of Chapter 9; the load's first-order autoregressive rows; the three readings and the discharge reading at each step; and the priors. The information is taken at the record's truth, seed 51, which makes each row of the table below a size and not an estimate. **This chapter computes the size and stops. Chapter 24 owns the estimator.**

**Building blocks.** The Jacobian of the stack, by the plant's own integration; the covariance $(\mathbf{J}^\top\mathbf{J})^{-1}$; the gradient of $T^*$ on $(g, g_w, b)$.

**The plant's numbers.** Chapter 5's zero row is a statement about steady window means, and this asks about the rows in time. Table 10.8 gives the sizes. Each row is a set of rows; the columns give the spreads of $g_w$ and $b$ and of $T^*$ at 45 kW, $r = 18°\mathrm{C}$. The static rows treat a window as Chapter 6 does, and "as if steady" counts every step of a later stretch as a steady row at that stretch's condition.

Table 10.8: The information of rows in time against static rows, at the record's truth, seed 51.

| Rows | sd $g_w$ [–] | sd $b$ [K] | sd $T^*$ [K] |
|---|---|---|---|
| Static, 30 kW window | 0.05077 | 0.18683 | 0.1538 |
| Stacked, the same steady window | 0.05196 | 0.18686 | 0.1890 |
| Stacked, load step, steps 80–159 | 0.04611 | 0.17577 | 0.0779 |
| Stacked, steps 40–159 | 0.03849 | 0.14884 | 0.0744 |
| Static as if steady, $40 + 80$ steps | 0.04217 | 0.15578 | 0.0564 |
| Stacked, all steps 40–239 | 0.02462 | 0.08592 | 0.0439 |
| Static as if steady, $40 + 80 + 80$ with the stop | 0.02294 | 0.08041 | 0.0402 |

The steady window's jitter adds nothing to $b$, 0.18686 against 0.18683 K, which confirms Chapter 5 by another route. **What a step response adds on $b$ is its new operating points, not its transient:** the stack lies within 7 % of the as-if-steady rows, 4.5 % better on one and 6.9 % worse on the other. The stop counts, since at the stop the reading speaks again. With the discharge rows removed, sd $b$ is 0.20000 K in every window, the prior to the digits printed. In the stack the reading returns as a row because the stack conditions on the command, where the static graph used the integrator's row. That is Chapter 5's consumed reading, restated.

The stacked fit on the record's own readings, seed 51, all steps, gives $\hat{b} = 0.3453$ K with spread 0.0854 K and $T^*$ of 25.4163°C with spread 0.0438 K. That is one record's draw. Over 50 fresh records, seeds 10000 to 10049, $\hat{b}$ has mean 0.2330 K (standard error 0.0097 K), a spread across records of 0.0685 K and a stated spread of 0.0864 K; the truth lies within two stated standard deviations in 0.92 of them for $b$ and 0.90 for $g_w$. On the load step alone the mean is 0.0850 K with a stated spread of 0.1738 K, near the prior. Here the model class is right, the loads follow the record's process and $T_w$ is pinned.

## 10.4 The checks (step 5)

The checks test the update's two forms, the loop's simulator, the moment route and the Laplace normal behind the set point. Table 10.9 collects them.

### 10.4.1 Routes against each other

Each row of Table 10.9 is a check; the columns name it, give its two routes and report the gap against its stated tolerance.

Table 10.9: The checks of the chapter.

| Check | Routes | Gap against tolerance |
|---|---|---|
| The update on linear rows | rank-one form against precision form | at most $4.54 \times 10^{-12}$ of an sd in the mean; tolerance $10^{-9}$ |
| Order of the rows | four frozen rows, 24 orders, against batch | $2.28 \times 10^{-14}$ of an sd in the mean; tolerance $10^{-9}$ |
| The batch simulator | against the record itself, seed 51, 240 steps | 0.0 on $T$, $T_c$, $u$ and $c$; tolerance $10^{-9}$ |
| The linear map | against the plant's own spectral radius and critical gain at 30 kW nominal | 0.94670 both; 1.1478 both, to the digits printed |
| Moments of the map | against the records, steady window, all nine routes | within 0.0006 K on sd $T_c$ and 0.003 per hour on travel |
| The floor | frozen limit against a refit at $n = 10^5$ | 0.1023 K both, to the digits printed |
| The average release | 9-point Gauss–Hermite on the refit against the sure release | +0.0018 against 0.0018; +0.0395 against 0.0395; +0.0481 against 0.0482 K |
| The chance level | reweighted draws, $t_5$ proposal, seed 10, $N = 10^6$, effective size 602 789, at the state route's $r_{95} = 20.0135$°C after the record's 40 readings | $\Pr(T^* \le 27.5$°C) of 0.9533 (standard error 0.0004) against 0.9500 |

The moments and the records are compared like with like: the root of the mean within-window variance on both sides, 0.0339 against 0.0335 K on the raw loop. Table 10.6 reports the mean within-window spread, 0.0323 K, a different average of the same records. The chance level's gap is Chapter 6's non-normality again, on the safe side at this point.

### 10.4.2 The limit of the physics rows

The chapter works on the reduced state of Section 4.2.8. On the full state $(T, T_c, u, Q, g, g_w, b)$ with the record's forty further hall readings, the physics rows at a tolerance $\sigma_p$ and the valve's row at $0.01\,\sigma_p$, the fit gives $g_w$ of $-0.16432$, $-0.16753$ and $-0.16756$ and $b$ of 0.05607, 0.05740 and 0.05741 K at $\sigma_p = 0.1$, 0.01 and 0.001. The reduced state gives $-0.16756$ and 0.05741 K.

### 10.4.3 The answers against the stated truth

After the record's forty further hall readings the truth stands +1.43 standard deviations from the fit on $T^*$, $-1.40$ on $g$, +1.39 on $g_w$ and +1.28 on $b$. The same sign pattern as Chapter 6's reappears, because hall readings do not move the fit along that ridge.

## 10.5 The report (step 6)

The back-off's schedule is known in advance to 0.001 K. At the pinned gains and 30 kW, a filter with $L = 0.3$ halves the estimate's error, from 0.0998 to 0.0493 K, and cuts the phase margin from 60.7° to 39.2°; at 45 kW the same pair run from 0.1003 to 0.0484 K and from 66.9° to 48.1°. Retuned to the raw loop's phase margin, the same filter takes the scatter down 54.5 % and the load-step peak up ×2.50, so a filter is a trade among travel, scatter and peak, and no route dominates. What stands at every gain is that the PI row has no covariance in it. The sensor's offset leaves the RMS of the true air about its set point at 0.3023 K raw and 0.3030 K at $L = 0.3$, against 0.0422 K with a reference. All percentages below are this example's.

### 10.5.1 The answers and their bands

Table 10.10 gives the answers in the order of Table 10.1. Each row is a question; the columns give the answer, what it rests on or its band, and the stated truth where one applies.

Table 10.10: The report for one more reading on the closed-loop hall. State-route values, with the fresh records' seeds in Section 10.3.5.

| Control engineer's question | Answer | Band or condition | Stated truth |
|---|---|---|---|
| What does one reading do to the belief? | The band narrows by an amount known in advance; the mean moves with the innovation. A reference takes sd $b$ from 0.1899 to 0.0484 K; a hall reading leaves it | route A against the refit: at most 0.0058 K on sd $T^*$, mostly the moved gradient (frozen, $-0.0006$ K) | $b = 0.30$ K |
| What is known before the reading comes? | Back-off 0.2565, 0.2170, 0.1893 K at $n = 0$, 40, 160 | covariance's dependence on the value at most 0.0009 K of back-off; the mean is not known; at $n = 40$ the stated 0.95 covered 0.820 of 300 fresh records | – |
| What can more hall readings not remove? | At one load, a back-off of 0.1683 K; twenty minutes takes 0.0395 K of the 0.0882 K removable. Across two loads the floor goes | two loads: the same path, a true hall sensor, metered heat. Prior on $b$: floor 0.0752 to 0.3772 K | $T^* = 25.4586^\circ$C |
| What does one reading do to the window? | $r_{95}$ gains the sure release on average: 0.0018, 0.0395, 0.0482 K. Reading by reading it moves either way | shares lowering it: 0.4613, 0.3173, 0.2982; 19.8714°C after a reference at the truth | 20.0414°C |
| Should the PI row act on the reading or on a filtered estimate? | At the pinned gains, on the reading for the air's sake. At $L = 0.3$ and 30 kW: estimate error at 0.495 of raw, travel $-41.7$ %, scatter $+11.3$ %, peak $\times 1.305$, margins 4.65 and 39.2°. Retuned to raw's phase margin: scatter $-54.5$ %, peak $\times 2.50$; a trade, with no route dominating | 1000 records, seeds 10000 to 10999. The row has no $\mathbf{\Sigma}$. A filter that knows the move halves the effective gain; raw at its crossover matches it except on travel | mean $T_c - r = -0.300$ K on every filter |
| What do a step response's rows add? | New operating points. Stacked, all steps, sd $b$ 0.0859 K; $\hat{b} = 0.3453$ K is one record's draw | 50 records, seeds 10000 to 10049: mean $\hat{b}$ 0.2330 K (se 0.0097). The estimator is Chapter 24's | $b = 0.30$ K |

### 10.5.2 The entry in the commissioning log

No filter of the reading on this loop at the pinned gains for the air's sake; the filter that knows the move lowers the scatter only as a lower gain would, and is entered as a detune if it is taken. If valve wear is the reason for a filter, the entry says so and prices it: at $L = 0.3$, 41.7 % less travel for a phase margin of 39.2° and a load-step peak of 0.2430 K. If the gains are retuned with it, the entry states the trade: at $K_p = 0.0243$ a scatter of 0.0147 K and a peak of 0.4659 K, against 0.0323 and 0.1863 K. The filter's gain is not to be taken from the quiet loop. The back-off's schedule is entered in advance: 0.2170 K after twenty minutes of hall readings, on a floor of 0.1683 K at this load, with the mean to be read when the readings come. Its stated 95 % covered 0.820 of 300 fresh records, so the entry does not call it a 95 % margin outright. The offset of the discharge sensor stays open, and it is not touched by any filter. The verdict rests on the stated truth's fresh records at these gains and on one record's draw of the belief.

### 10.5.3 Common mistakes

Table 10.11 lists five mistakes this chapter's plant makes easy. Each row names the mistake, its form on this hall and the correction.

Table 10.11: Five mistakes in reading a new reading on a loop.

| Mistake | On this hall | Correction |
|---|---|---|
| Taking a better estimate for a better loop | error 0.0998 to 0.0493 K with scatter 0.0323 to 0.0360 K and phase margin 60.7° to 39.2° at $L = 0.3$ | score the true quantity and the margin, on the closed loop |
| Saying a fixed law runs on the posterior | the PI row (10.2) has one estimate in it and no $\mathbf{\Sigma}$ | the row does not exercise coupling 4; recomputing $r$ from the back-off would, by the limit route (Chapter 16), and the loop study pins $r$; say which coupling a controller keeps |
| Taking a filter's gain from the quiet loop | 0.0118 K a step gives $L = 0.1108$, and near it a phase margin of 17.3° | the quiet is the loop's doing; a random walk fitted inside the loop does not describe the reading's source |
| Writing a sensor's readings as independent rows | sd $b$ 0.0348 claimed at $n = 40$, 0.1379 K true; a back-off of 0.2065 against 0.2334 K | carry the shared offset as an entry |
| Reading a floor as a property of the sensor | 0.1683 K at one load; sd $b$ 0.0006 K with two loads | state the condition: one load; two loads read it through the physics if the path is the same, the hall sensor true and the heat metered |

## 10.6 The next measurement (step 7)

The loop acts on two things that a reading might narrow: the supply air $T_c = r - b$, spread 0.1899 K, and through the set point's margin $T^*$, spread 0.1559 K. The ranking below is the frozen preposterior, so it asks only of the band.

### 10.6.1 The candidates

Table 10.12 ranks the candidates by equation (4.73). Each row is a candidate, the permanent sensor among them; the columns give the spread of $T_c$, the spread of $T^*$ at 45 kW, $r = 18$°C, and the back-off after it.

Table 10.12: The candidates, on frozen rows at $\hat{\mathbf{z}}$. The two rows that change a condition add a load entry and rest on the same path and, for the set-point move, on the pinned $T_w$ and the coil's law; their rows assume a settled hall, and the last column is the steps until its mean stands within 0.02 K, on the raw loop.

| Candidate | sd $T_c$ [K] | sd $T^*$ [K] | Back-off [K] | Settling [steps] |
|---|---|---|---|---|
| Discharge reading, any number | 0.1899 | 0.1559 | 0.2565 | 0 |
| One hall reading | 0.1899 | 0.1549 | 0.2547 | 0 |
| One reference discharge reading, 0.05 K | 0.0484 | 0.1266 | 0.2083 | 0 |
| Permanent second discharge sensor, 40 readings | 0.1379 | 0.1419 | 0.2334 | 0 |
| Hall, 40 readings, or a second hall sensor | 0.1899 | 0.1319 | 0.2170 | 0 |
| Hall, valve, load meter, 40 steps | 0.1899 | 0.1289 | 0.2121 | 0 |
| 40 steps at 45 kW, the same sensors | 0.1591 | 0.0733 | 0.1206 | 89 |
| 40 steps at 30 kW and $r = 20°$C, the same sensors | 0.0901 | 0.0987 | 0.1623 | 91 |

### 10.6.2 The recommendation

For one reading, the reference does the most on both quantities. With no new instrument, 40 steps at 45 kW does the most for $T^*$, and **a set-point move to** 20°C **does the most for** $T_c$**, second only to the reference.** Neither is twenty minutes. Their rows assume a settled hall, so with settling they take 129 and 131 steps, 64.5 and 65.5 min, and Chapter 9's rows in time show what readings taken during the move are worth. A move teaches. This is coupling 3 of Section 7.1.3, stated here and owned by Chapters 22 to 24. It leans on $T_w$ pinned and on the coil's law, and the record has no such window to test it. As a permanent instrument, a second hall sensor gives the band more than a second discharge sensor, and a hall sensor's own offset is then the entry to guard, by the what-if of Section 10.3.3. The filter is not a candidate, since it adds no row.

### 10.6.3 What the analysis showed

The sheet offered more data as one answer to two requests, and the analysis separates them. A further hall reading narrows the band by an amount known in advance and moves the mean by an amount that is not, and the band has a floor that is a floor at one load only. A filter is not more data: it adds no row to the belief and changes the plant through the controller's row. It did what a filter does, and at the pinned gains the loop was worse for it, because the noise that moves the air lies where the filter removes little and adds lag; retuned, it traded travel, scatter and peak, with no route dominating. The row never asked for the band: recomputing $r$ from the back-off would exercise that coupling, and the loop study pinned $r$. A step response's value lay in the operating points it visited. What the engineer takes back is that a better estimate and a better loop are two claims, and each has to be scored against the air.

### 10.6.4 Limits of the model

**The Laplace prior after a large move.** Route B is not recommended, and route A's error grows with the reading's pull on the earlier rows, as the reference showed. Only route C reports a value after a large move.

**Independent rows.** Rows are taken as independent over 30 s. Load wander is not studied beyond Chapter 5's amendment, and it would slow the schedule.

**The second load.** The path is taken as the same at both loads, and $T_w$, the valve and the load meter as without offsets, with no unmetered heat. A path 1 % better at 45 kW shifts the intercept 0.1536 K, and 1 kW unmetered shifts it 0.1724 K. A hall sensor with an offset reads only $b - d$; a prior of 0.10 K on $d$ leaves sd $b$ at 0.0886 K. The plant's path is constant by construction, so the record cannot test this, and the rows assume a settled hall.

**The fixed row.** The PI row is fixed and not re-tuned; Chapter 14 re-tunes. The filter's verdict holds at $K_p = 0.10$, $K_i = 0.002$, with the grid and the retuned rows above as all that is claimed at other gains.

**Fresh records from the stated truth.** The loop study draws from the truth, so model error is not tested, and the margins are the truth's and not carried under the band.

**The stated rules and the example's percentages.** The 0.001 K rule is a stated choice. The percentages are this example's and not properties of control or of filters.

**The constructed record.** The record is constructed from a stated truth, and the reference reading at 17.70°C was constructed for the illustration. Chapter 11 asks in general which entries the rows of a loop can reach.

## 10.7 What to retain

**Definition 10.1** (Schedule of the back-off)**.** The schedule of the back-off is the back-off of a threshold at a scenario node as a function of the number of further readings,

$$\Phi^{-1}(1-\alpha)\sqrt{\mathbf{g}^{\mathsf{T}}\left(\boldsymbol{\Sigma}^{-1} + n\mathbf{h}\mathbf{h}^{\mathsf{T}}/\sigma^2\right)^{-1}\mathbf{g}}.$$

On linear rows it is known before any reading; on a nonlinear row it holds to a tolerance that has to be measured. Its limit is the floor, which belongs to the row's line and so to the condition at which that row is read.

Three conclusions carry beyond this hall, as statements of form and not of size. A scalar row moves the mean by the innovation and the covariance by a known outer product, so the band can be planned and the mean cannot. A floor set by a shared offset belongs to the lines being read, and a line at a new condition can break it if the physics between the conditions holds. And a fixed law takes a point estimate, with no covariance in its row: its estimate's quality and the loop's quality are different questions, to be scored separately on the closed loop, and what a filter costs depends on the gains it is retuned to, so it is a trade and not a verdict.

Table 10.13 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 10.13: The terms of the chapter and where each is defined. Units for the closed-loop hall.

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| PI row on an estimate | eq. (10.2) | – | a point estimate, no $\mathbf{\Sigma}$ | 10.1.4 |
| Scalar filter | eq. (10.3) | K | a fit of the reading, gain $L$ | 10.1.4 |
| Rank-one update | eq. (10.6) | mixed | one scalar row's move of the belief | 10.2.2 |
| Routes A, B, C | frozen; Laplace prior; refit, eq. (10.7) | – | three ways with a nonlinear row | 10.2.3 |
| Schedule of the back-off | eq. (10.8) | K | the band against the count | 10.3.2 |
| Floor | eq. (10.9) | K | the row's value known | 10.3.3 |
| The line in the load | eq. (10.10) | K | two loads read the intercept | 10.3.3 |
| Sure release | first term of eq. (10.11) | K | what a reading gives on average | 10.3.4 |
| Share lowering $r_{95}$ | eq. (10.12) | – | readings for which the set point falls | 10.3.4 |

## 10.8 Exercises

**Exercise 10.1.** *Formulation.* A level sensor of standard deviation $\sigma_\ell$ is added on a tank whose level $\ell$ is an entry of the state. Write its new row and name what its covariance update needs. (The row is $(\ell - y_\ell)/\sigma_\ell$, with $\mathbf{h}$ the unit vector of the entry $\ell$ when it is a free entry, and otherwise the gradient of the level at $\hat{\mathbf{z}}$. Equation (10.6) then needs $\mathbf{\Sigma}$, $\mathbf{h}$ and $\sigma_\ell$. It needs no reading. The mean's update alone needs $y_\ell$.)

**Exercise 10.2.** *Derivation.* Reach equation (4.74) from the precision form. (Let $\mathbf{\Sigma}'^{-1} = \mathbf{\Sigma}^{-1} + \mathbf{h}\mathbf{h}^\mathsf{T}/\sigma^2$. The candidate $\mathbf{\Sigma}' = \mathbf{\Sigma} - \mathbf{\Sigma}\mathbf{h}\mathbf{h}^\mathsf{T}\mathbf{\Sigma}/S$, with $S = \mathbf{h}^\mathsf{T}\mathbf{\Sigma}\mathbf{h} + \sigma^2$, multiplied by it gives $\mathbf{I} + \mathbf{\Sigma}\mathbf{h}\mathbf{h}^\mathsf{T}(1/\sigma^2 - 1/S - \mathbf{h}^\mathsf{T}\mathbf{\Sigma}\mathbf{h}/(S\sigma^2)) = \mathbf{I}$, since the bracket is $(S - \sigma^2 - \mathbf{h}^\mathsf{T}\mathbf{\Sigma}\mathbf{h})/(S\sigma^2) = 0$. With $\mathbf{K} = \mathbf{\Sigma}\mathbf{h}/S$, $\mathbf{\Sigma}' = \mathbf{\Sigma} - S\mathbf{K}\mathbf{K}^\mathsf{T}$. The mean then follows by equation (10.5).)

**Exercise 10.3.** *Derivation.* Show why the noise-to-air gain is flat below the crossover and give its peak on the raw loop at 30 kW. (With $P$ the plant from valve to air and $C$ the PI row, the air's response to the sensor noise is $-PC/(1+PC)$. Well below the crossover $|PC| \gg 1$ and the magnitude tends to one; well above it $|PC| \ll 1$ and the magnitude tends to $|PC|$, which falls. A gain of one means that the noise is followed, and a filter that leaves the low frequencies alone cannot change that. The peak sits near the crossover of 0.00718 rad/s: 1.076 at 0.00342 rad/s.)

**Exercise 10.4.** *Application.* Find the back-off at 45 kW after $n = 40$ further hall readings. (The rows act as one reading of variance $0.5^2/40$, with predictive standard deviation 0.1117 K. The mean of $T^*$ moves $c = +0.7442$ per kelvin of innovation, so the move has spread $0.7442 \times 0.1117 = 0.0832$ K. The spread after is $\sqrt{0.1559^2 - 0.0832^2} = 0.1319$ K, and the back-off is $1.6449 \times 0.1319 = 0.2170$ K.)

**Exercise 10.5.** *Application.* Give the filter's gain for a walk of step 0.05 K on the reading, and say why it says nothing of the loop. (Equation (10.3) with $\sigma_w/\sigma_v = 0.5$: $L^2/(1-L) = 0.25$, so $L = 0.3904$. It is the gain of a fit on a random walk; the loop's margins and the air's scatter are not in it, and Section 10.3.5 measures those separately.)

**Exercise 10.6.** *Application.* From Table 10.4, say what writing the second discharge sensor's readings as independent rows claims at $n = 40$, and what is true. (It claims a spread of $b$ of 0.0348 K and a back-off of 0.2065 K. With the offset carried as an entry they are 0.1379 and 0.2334 K.)

# Chapter 11

# Observability and Sensor Loss

## 11.1 A loop that holds a reading it may lose

The commissioning log at 23:00 shows the chilled-water loop of the two-unit hall at rest. The return reading stands on its set point, a loop error of 0.00 K, with the supply at 14.00°C. The commissioning sheet, at rated units, survey fractions and 200 kW, puts the rack-3 inlet at 23.66°C, a margin of 3.34 K to its limit of 27°C. A work order now takes the return sensor out. The engineer concludes that the loop must go to manual at its last supply temperature, and that 3.34 K covers whatever the missing sensor might have said.

The sibling volume *Operations and Control by Derivation*, in its Chapter 17, asks what a plant's band loses when a sensor drops. There nothing moves the chilled water in answer to the reading, so a lost row only widens a fit. Here a controller moves the supply to hold that reading, so the same loss also removes a row of the plant, and the two ways of carrying on are different plants. The sheet does not say which of them holds the limit, or whether the other sensors could stand in. The question of this chapter is

*what the sensors can resolve of what the controller needs, and what the loop does when they cannot.*

**Jargon: observability for a question, unobservable entry, dropout, stuck reading, frozen actuator, virtual reading, air-balance route, virtual-reading loop, loop-dependent reading, down-date, refit, belief aggregation, aggregate map, marginal by Schur complement, void node**

Chapter 6 named *observability*: whether the readings determine an entry. *Observability for a question* asks whether they determine what a named quantity needs of that entry, which is read from the entry's gradient times its spread. A *dropout* is a reading that is gone, and a *stuck reading* is one that stays connected and stops changing. A *frozen actuator* holds its last position with nothing correcting it. A *virtual reading* is a quantity computed from other sensors and used in place of one; the *air-balance route* is the virtual reading of the return built from the discharges and the load meter, and a *virtual-reading loop* is a controller closed on such a reading. A *loop-dependent reading* is one a running loop's reading is computed from, whose loss sends the loop to another route with a wider band. It is not a *critical measurement* in the sense of state estimation, whose loss leaves the state unobservable: each loss here is absorbed by another route. A *down-date* removes a row's information from the fit at its MAP state and leaves the state where it is; a *refit* removes the row and solves again, so the state moves too. *Belief aggregation* passes a parent level a smaller description of a posterior: an *aggregate map* $\mathbf{a}(\mathbf{z})$, and its *marginal by Schur complement*. A *void node* is a closed node whose actuator would lie past a stop: its formula describes no plant.

**Key idea**

Observability is relative to a question, and a loop changes the question. Closing on the return takes the coils out of the rack inlet's gradient, so the readings that served the coils stop mattering and an entry no reading touches sets the band. When the reading goes, the two plants that remain, with the airflow pinned, answer a moving load with opposite signs, so each breaks its limit on its own side.

### 11.1.1 Observability read against a need

Two objects decide whether a sensor set is enough, and they answer different questions. The first is the column of an entry in the Jacobian of the reading rows, together with the ratio of the entry's posterior spread to its prior spread. A zero column, or a ratio of one, says the readings hold nothing on that entry. The second is the gradient of the quantity the controller needs. By the delta method (4.53) and the budget (4.54),

$$\mathrm{Var}(f \mid \mathbf{y}) \approx \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g} = \sum_i \left(\mathbf{g}_i\sqrt{\Sigma_{ii}}\right)^2 + \sum_{i\neq j}\mathbf{g}_i\mathbf{g}_j\Sigma_{ij}, \tag{11.1}$$

and the entry's term $\mathbf{g}_i\sqrt{\Sigma_{ii}}$ is what it contributes to the band of $f$. The two objects together give four classes: an entry may be seen or unseen, and it may be needed or not. An unseen entry that nothing needs is harmless. A seen entry that is needed is the case a sensor list is usually drawn for. An unseen entry that is needed is the band's floor, and no rearrangement of the existing sensors reaches it. Which class an entry is in depends on $f$, and a loop changes $f$.

### 11.1.2 The two-unit hall with its loop (step 0)

This chapter's plant is larger than the one-path hall of Parts I and II, because the question needs more than one coil and more than one place to measure. Figure 11.1 draws it. Its upper part is the hall. Chilled water at the supply temperature $T_s$ feeds two coils. Return air at $T_r$ passes through each coil, leaves at a discharge temperature $T_{d1}$ or $T_{d2}$, and the two streams mix into the cold aisle at $T_c$. Three racks take air from the cold aisle, each recirculating a fraction $\beta_k$ of the return, so that its inlet is $T_k$. The air leaves the racks to the return, carrying the load $Q$ and a pinned other heat $Q_o$. The lower part of the figure repeats the essentials three times, for the three graphs of this chapter. Green boxes hold free quantities, gray boxes pinned ones, blue boxes quantities the rows fix, and dashed arrows lead to readings.

**The physical rows.** With $G_u = G_{\mathrm{rated},u}\mathrm{e}^{g_u}$, $G_{\mathrm{rated},u} = 7\,\mathrm{kW/K}$, the airflow capacities $C_u = 10\,\mathrm{kW/K}$ pinned, and $Q_o = 8\,\mathrm{kW}$ pinned, the hall has seven physical rows, each the limit $\sigma_p \to 0$ of a row with its tolerance:

$$\begin{aligned}
(G_1+G_2)(T_r - T_s) &= Q + Q_o && && \text{heat balance of the coils,}\\
C_u\,(T_r - T_{d,u}) &= G_u\,(T_r - T_s), && u = 1,2 && \text{each coil's air side,}\\
(C_1+C_2)\,T_c &= C_1T_{d1} + C_2T_{d2} && && \text{mixing,}\\
T_k &= T_c + \beta_k\,(T_r - T_c), && k = 1,2,3 && \text{each rack's inlet.}
\end{aligned} \tag{11.2}$$

Open, with $T_s$ pinned, the graph has 13 entries, $(Q, g_1, g_2, \beta_1, \beta_2, \beta_3)$ and the seven temperatures $T_r, T_{d1}, T_{d2}, T_c, T_1, T_2, T_3$, on 7 rows, which leaves 6. Closed by Chapter 5's construction, the supply is free, the reading's offset $b_r$ is an entry, and the controller's row is added,

$$T_r + b_r = r, \tag{11.3}$$

so 15 entries on 8 rows leave 7. The integrator adds one entry and one row, net zero, as it did there.

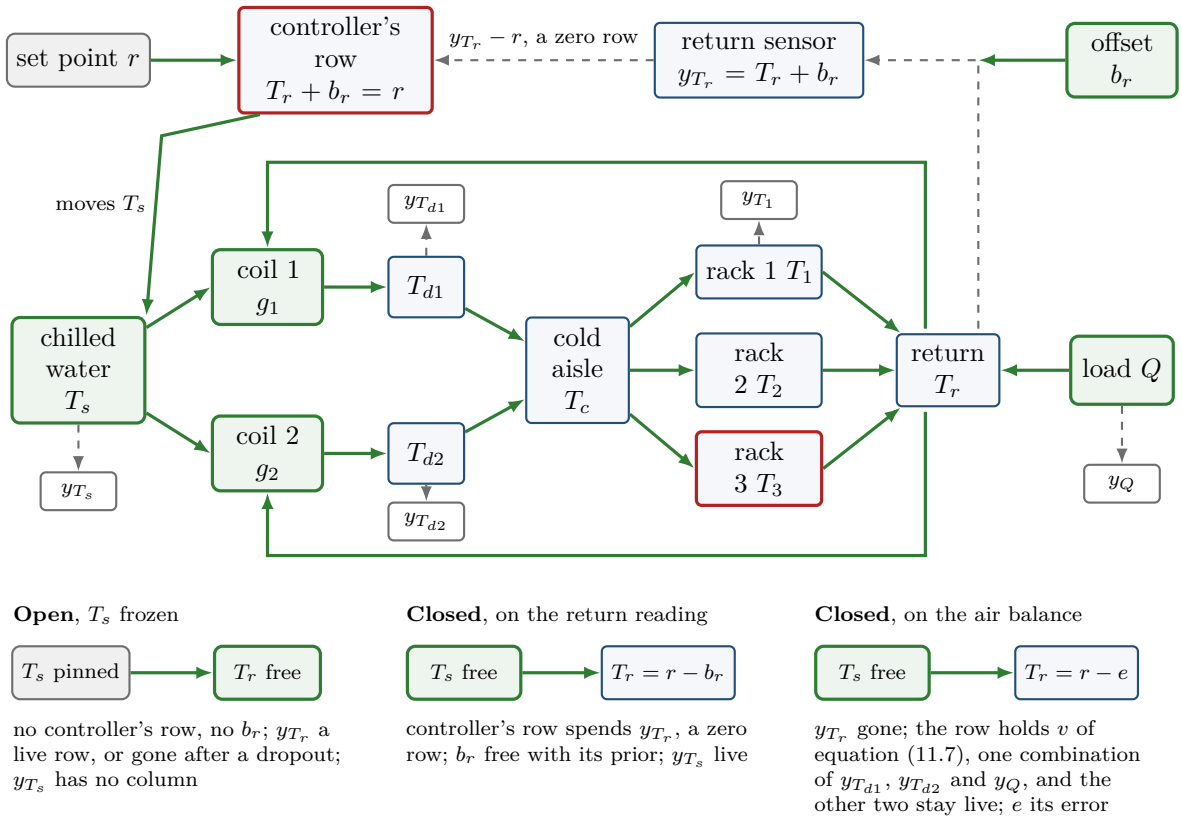



Figure 11.1: The two-unit hall with its loop, and the three graphs of the chapter. Free (green): the load, the coils' log-derates, the supply, the offset. Pinned (gray): the set point; the other heat $Q_o$ and the airflows are pinned and not drawn. Blue: quantities the rows fix, among them the rack inlets, with $\beta_k$ a free entry carried on each rack. The protected inlet $T_3$ (red) has no sensor. Dashed: readings. The return sensor's reading goes into the controller's row (red), and what is left of it is the loop's error, which holds no entry.

**The reduced state.** The free entries, on the reduced state of Section 4.2.8, are

$$\mathbf{z} = (Q,\ g_1,\ g_2,\ \beta_1,\ \beta_2,\ \beta_3,\ b_r)^{\mathsf{T}}, \tag{11.4}$$

in kW, K and dimensionless entries, and with $\sum G = G_1 + G_2$ and $\sum C = C_1 + C_2 = 20\,\mathrm{kW/K}$ the rows give the dependent quantities,

$$T_r = r - b_r, \quad T_s = T_r - \frac{Q + Q_o}{\sum G}, \quad T_{d,u} = T_r - \frac{G_u}{C_u}\frac{Q + Q_o}{\sum G}, \quad T_c = T_r - \frac{Q + Q_o}{\sum C}. \tag{11.5}$$

The last relation is the air balance: the two coils together take out $Q + Q_o$, whatever their split, and $\sum C$ turns that into a temperature drop in the air. It follows that every rack inlet is

$$T_k = r - b_r - (1 - \beta_k)\frac{Q + Q_o}{\sum C}, \tag{11.6}$$

and this line carries much of the chapter. The coils are not in it.

**Modeling choices at the graph.** Five constants are chosen here. *The set point* is $r = 30.33°\mathrm{C}$, the window's own return reading. This is a stated choice: it keeps the plant at $T_s = 14°\mathrm{C}$ and the record's readings unchanged, and it would have to be restated for a loop commissioned elsewhere. *The offset's prior* is $b_r \sim \mathcal{N}(0,\ 0.30^2)$ in K, the reading's only stated tolerance. It buys a state the fit can solve, and it is tested in Section 11.4.4. *The supply sensor* has a standard deviation of 0.05 K, chosen because the plant's description calls it precise without a value. *The supply's stops* are 8 and 20°C. A closed node past one is void. *A node's reading error is a fresh draw*, as the record draws noise per window, so its mean at the node is zero; this is what lets a node's band carry the reading's tolerance again, and it fails for an error that persists, which moves the node's mean as well as its band, as Section 11.3.4 tests. The chapter reads one steady window, and capacity is left out; Limits returns to that.

**The virtual reading.** Rearranging the air balance gives a value of the return from sensors other than the return's,

$$v = \tfrac{1}{2}(y_{T_{d1}} + y_{T_{d2}}) + \frac{y_Q + Q_o}{\sum C}, \tag{11.7}$$

with equal airflows, needing no coil and no recirculation fraction.

**The scenario nodes.** Each what-if pins the load at $Q^*$ and says whether the loop is open or closed. A *frozen* node is open: the supply stays at the window's $T_s$, the function of the state in equation (11.5), so its belief is carried, and

$$T_3^*(\mathbf{z}; Q^*) = T_s(\mathbf{z}) + \frac{Q^* + Q_o}{\sum G} - (1 - \beta_3)\frac{Q^* + Q_o}{\sum C}. \tag{11.8}$$

A *hold* is the same node with $T_s$ pinned at an exact value, which is what a chosen supply set point would be. A *closed* node pins the set point and frees the supply. The controller's row then holds its reading, whose error $e$ at the node is a fresh draw of standard deviation $\sigma_e$, the sensor's 0.30 K or the virtual reading's own:

$$T_3^*(\mathbf{z}; Q^*, r) = r - e - (1 - \beta_3)\frac{Q^* + Q_o}{\sum C}, \qquad T_s^* = r - e - \frac{Q^* + Q_o}{\sum G}. \tag{11.9}$$

The closed form holds only while $8°\mathrm{C} \le T_s^* \le 20°\mathrm{C}$, which the chapter reads at the mean, with $e = 0$, and as a probability with $e$ in the spread. Past a stop the supply is pinned again, the set point is lost, and equation (11.9) describes no plant.

### 11.1.3 The control engineer's questions (step 1)

Table 11.1 lists the eight questions. The first column is the question in the engineer's words, the second the general form, the third its mathematical form on this hall with a star at a scenario node, the fourth its category, the fifth the computation and the last the subsection. The protected quantity is the rack-3 inlet $T_3$, unmeasured, against 27°C at 0.95.

Table 11.1: The questions asked of the two-unit hall and its loop. $\mathbf{y}$ is the five live readings and the six priors.

| Control engineer's question | General form | Mathematical form | Category | Computation | Answered in |
|---|---|---|---|---|---|
| What can these readings resolve? | moments, (2.10) | rank of the reading rows; sd against prior, per entry | estimation | posterior | 11.3.1 |
| What does the controller need of them? | moments of a derived quantity, (2.10) | $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ for $T_3$, the hold node and $T_s^*$ | estimation of a derived quantity | posterior | 11.3.2 |
| The loop's reading drops: what is the loop? | moments at a scenario node, (2.10) | $T_3^*$, $T_r^*$, $T_s$ frozen, open loop | prediction | posterior | 11.3.3 |
| Can a virtual reading replace it? | moments at a scenario node, (2.10) | closed on the air balance, $T_3^*$, closed loop | prediction | posterior | 11.3.4 |
| Hold or virtual loop, as the load moves? | chance constraint, (6.1) | $\Pr(T_3^* \le 27^\circ\mathrm{C} \mid \mathbf{y}) \ge 0.95$; $r_{95}$, $T_{s,95}$, crossing loads | prediction under a threshold | posterior | 11.3.5 |
| Which other reading matters most? | moments, row removed, (2.10) | the ranking by added band on $T_3$ | estimation | posterior | 11.3.6 |
| What does a parent level keep? | marginal of a derived quantity, (2.10) | $p(Q_{\mathrm{tot}}, G_{\mathrm{tot}} \mid \mathbf{y})$ by Schur complement | estimation of a derived quantity | posterior | 11.3.7 |
| Which sensor next? | preposterior variance, (4.73) | ranking by $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ against by entry spread | sensor placement | preposterior | 11.6.1 |

### 11.1.4 The evidence as rows (step 2)

**The readings and the priors.** The evidence is the 23:00 window: the supply $y_{T_s} = 14.00$°C (sd 0.05 K), the discharges $y_{T_{d1}} = 19.47$ and $y_{T_{d2}} = 20.11$°C (sd 0.30 K each), the load meter $y_Q = 193.04$ kW (sd 3 kW) and the rack-1 inlet $y_{T_1} = 22.18$°C (sd 0.30 K). Each is taken as an independent normal row given the state. The priors are $g_u \sim \mathcal{N}(0,\ 0.15^2)$ on the coils, from their rating, $\beta_k \sim \mathcal{N}(\text{survey},\ 0.08^2)$ with surveyed fractions 0.20, 0.35 and 0.50, and the offset's of step 0. The load has none, since the meter reads it.

**The reading the loop consumes.** The return reading $y_{T_r} = 30.33$°C is not a row of the closed graph. Its model is $T_r + b_r$, which equation (11.3) makes equal to $r$ whatever the state, so its line in the Jacobian is zero and its residual is the loop's error, as in Chapter 5. Five live readings then stand on seven entries in the closed graph. In the open graph, $y_{T_r}$ is a live row, $y_{T_s}$ has no column, and five readings stand on six entries.

**Which row acts on which entry.** By equations (11.5) and (11.6), the supply and the two discharges depend on $Q$, both coils and $b_r$; the meter on $Q$ alone; the rack-1 inlet on $Q$, $\beta_1$ and

$b_r$. With $*$ for an entry that can be nonzero and the columns of equation (11.4), the $11 \times 7$ Jacobian has the pattern

$$\mathbf{J} = \begin{bmatrix} * & * & * & 0 & 0 & 0 & * \\ * & * & * & 0 & 0 & 0 & * \\ * & * & * & 0 & 0 & 0 & * \\ * & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & * & 0 & 0 & * \\ 0 & * & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} y_{T_s} \\ y_{T_{d1}} \\ y_{T_{d2}} \\ y_Q \\ y_{T_1} \\ \text{prior on } g_1 \\ \text{prior on } g_2 \\ \text{prior on } \beta_1 \\ \text{prior on } \beta_2 \\ \text{prior on } \beta_3 \\ \text{prior on } b_r \end{matrix} \tag{11.10}$$

Two facts are read here before any number. The columns of $\beta_2$ and $\beta_3$ hold only their priors, since no reading stands at those racks. The five reading rows stand on five columns besides those, so they leave no spare row, and that matters in step 3.

## 11.2 The posterior of the state (step 3)

All eight questions read one posterior, of the reduced state (11.4) under the eleven rows of step 2. The open graph's fit is computed beside it, because the comparison is the chapter's first result.

### 11.2.1 The rows, the normal equations and the covariance

Stacked in the order of the pattern (11.10), each divided by its standard deviation, the rows form

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T_s(\mathbf{z}) - 14.00)/0.05 \\ (T_{d1}(\mathbf{z}) - 19.47)/0.30 \\ (T_{d2}(\mathbf{z}) - 20.11)/0.30 \\ (Q - 193.04)/3 \\ (T_1(\mathbf{z}) - 22.18)/0.30 \\ g_1/0.15 \\ g_2/0.15 \\ (\beta_1 - 0.20)/0.08 \\ (\beta_2 - 0.35)/0.08 \\ (\beta_3 - 0.50)/0.08 \\ b_r/0.30 \end{bmatrix}, \tag{11.11}$$

with $\mathcal{L}(\mathbf{z}) = \frac{1}{2}\,\mathbf{r}^{\mathsf{T}}\mathbf{r}$ and the dependent quantities from equations (11.5) and (11.6). Each iteration solves

$$(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \qquad \mathbf{z}^{(k+1)} = \mathbf{z}^{(k)} + \Delta\mathbf{z}, \tag{11.12}$$

on these normalized rows, whose information is $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$; with unnormalized rows it is $\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$. The rows are nonlinear through $\mathrm{e}^{g_u}$ and the products with $\beta_1$, so the curvature term of equation (4.44) is kept,

$$\mathbf{\Sigma} = \left(\mathbf{J}^{\mathsf{T}}\mathbf{J} + \mathbf{C}\right)^{-1}\Big|_{\hat{\mathbf{z}}}, \tag{11.13}$$

and $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$.

### 11.2.2 The fit of the window

**The iterations.** Table 11.2 runs equation (11.12) from the meter's load, rated coils, survey fractions and no offset. Each row is an iteration; the columns give the entries that move, the objective and the norm of its gradient. Two steps bring the objective to its minimum at the digits printed, and $\beta_2$ and $\beta_3$ do not move from 0.35 and 0.50.

Table 11.2: Gauss–Newton iterations of the closed graph, equation (11.12), on the reduced state.

| $k$ [–] | $Q$ [kW] | $g_1$ [–] | $g_2$ [–] | $\beta_1$ [–] | $b_r$ [K] | $\mathcal{L}$ [–] | $\lVert \mathbf{J}^{\mathsf{T}}\mathbf{r} \rVert$ [–] |
|---|---|---|---|---|---|---|---|
| 0 | 193.0400 | 0.0000 | 0.0000 | 0.2000 | 0.0000 | 780.0291 | $8.04 \times 10^{3}$ |
| 1 | 194.6168 | −0.0789 | −0.1401 | 0.2208 | 0.2834 | 3.1061 | $4.34 \times 10^{2}$ |
| 2 | 194.6283 | −0.0735 | −0.1341 | 0.2204 | 0.2801 | 1.3181 | $1.20 \times 10^{0}$ |
| 3 | 194.6273 | −0.0735 | −0.1340 | 0.2204 | 0.2802 | 1.3181 | $1.79 \times 10^{-4}$ |
| 4 | 194.6273 | −0.0735 | −0.1340 | 0.2204 | 0.2802 | 1.3181 | $1.90 \times 10^{-7}$ |

**The Jacobian at the MAP state.** In the columns of equation (11.4) and the rows of equation (11.11),

$$\mathbf{J}(\hat{\mathbf{z}}) = \begin{bmatrix} -1.58403 & 165.34313 & 155.62391 & 0 & 0 & 0 & -20 \\ -0.17171 & -16.87013 & 16.87013 & 0 & 0 & 0 & -3.33333 \\ -0.16162 & 16.87013 & -16.87013 & 0 & 0 & 0 & -3.33333 \\ 0.33333 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.12993 & 0 & 0 & 33.77122 & 0 & 0 & -3.33333 \\ 0 & 6.66667 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6.66667 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 12.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 12.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 12.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3.33333 \end{bmatrix}. \tag{11.14}$$

The zeros are the pattern's. The first three rows show how the coils are seen. The supply reading carries the sum of the coils, with large equal-signed entries; the two discharges carry their difference, with opposite signs.

**The covariance.** The Laplace covariance (11.13), in the same order and in mixed units, is

$$\boldsymbol{\Sigma} = 10^{-4} \begin{bmatrix} 76590 & 205.6 & 226.1 & 34.05 & 0 & 0 & -2604 \\ 205.6 & 5.719 & -2.394 & 1.638 & 0 & 0 & 10.83 \\ 226.1 & -2.394 & 6.339 & 1.644 & 0 & 0 & 10.09 \\ 34.05 & 1.638 & 1.644 & 9.881 & 0 & 0 & 23.66 \\ 0 & 0 & 0 & 0 & 64.00 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 64.00 & 0 \\ -2604 & 10.83 & 10.09 & 23.66 & 0 & 0 & 374.4 \end{bmatrix}.$$

The correlations above 0.3 in size are $Q$ with $g_1$ at 0.311 and with $g_2$ at 0.325, $Q$ with $b_r$ at $-0.486$, $g_1$ with $g_2$ at $-0.398$ and $\beta_1$ with $b_r$ at 0.389. The two zero rows are $\beta_2$ and $\beta_3$, which share nothing with any other entry, as the pattern required.

Table 11.3: The MAP state of the same window in the two graphs, with $\sqrt{\Sigma_{ii}}$ in parentheses.

| Entry [unit] | Closed, $T_s$ free, $y_{T_s}$ a row | Open, $T_s$ pinned, $y_{T_r}$ a row |
|---|---|---|
| $Q$ [kW] | 194.6273 (2.7676) | 194.6275 (2.7675) |
| $g_1$ [–] | −0.0735 (0.0239) | −0.0736 (0.0237) |
| $g_2$ [–] | −0.1340 (0.0252) | −0.1341 (0.0250) |
| $\beta_1$ [–] | 0.2204 (0.0314) | 0.2204 (0.0314) |
| $\beta_2$ [–] | 0.3500 (0.0800) | 0.3500 (0.0800) |
| $\beta_3$ [–] | 0.5000 (0.0800) | 0.5000 (0.0800) |
| $b_r$ [K] | 0.2802 (0.1935) | no such entry |

**The MAP state.** Table 11.3 sets the closed fit beside the open fit of the same window, $T_s$ pinned at 14°C and $y_{T_r}$ a row. Each row is an entry; the columns give the value and spread in each graph.

The two columns differ only in the fourth decimal place, and Section 11.2.3 gives the reason. The misfit of the closed fit is $2\mathcal{L}(\hat{\mathbf{z}}) = 2.6361$ on $11 - 7 = 4$ degrees of freedom, $P = 0.6204$, of which the readings hold 0.6601 and the priors 1.9761. The open fit gives 2.6370 on $10 - 6 = 4$, $P = 0.6203$. The loop's error of 0.00 K enters neither.

### 11.2.3 A prior at the reading's tolerance is the reading's row

The open graph has the row $(T_r - y_{T_r})/0.30$ with $T_r$ free. The closed graph has the prior row $b_r/0.30$, and by the controller's row (11.3) $b_r = r - T_r$. With $r = y_{T_r}$, the choice of step 0,

$$\frac{b_r}{0.30} = \frac{y_{T_r} - T_r}{0.30} = -\frac{T_r - y_{T_r}}{0.30}, \tag{11.15}$$

the same row up to its sign, which the objective does not read. The other difference between the graphs is the supply, free with a row against pinned, and a row at a vanishing standard deviation is a pin. So the two graphs should coincide as the supply reading's tolerance goes to zero. With $y_{T_s} = 14$ at sd $10^{-4}$, the closed fit against the open fit has a largest entry difference of $3.53 \times 10^{-10}$, a largest spread difference of $3.46 \times 10^{-8}$ of the spread, and $2\mathcal{L}$ of 2.6370 against 2.6370, each within a stated tolerance of $10^{-3}$. At the stated 0.05 K the difference is the fourth decimal place of Table 11.3. The identity is that of a prior at the reading's tolerance, 0.30 K; at another width the row is a different row, and the closed fit's means move (Section 11.4.4).

> **Principle 11.1.** *Closing a loop on a reading turns the reading's error into a plant offset, and the prior on that offset at the sensor's tolerance is the reading's row. The information has not moved. At the window the loop changes nothing in the band; it changes which entries the controller needs, and everything at a node.*

## 11.3 The answers (step 4)

Every answer reads $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 11.2.2 by the state route, which is the value reported. The routes that a sensor's loss brings in are named where they enter.

### 11.3.1 What the readings resolve

**Control engineer's question.** What can these readings resolve?

**General form.** Moments of an entry, equation (2.10), under the posterior.

**Specialized form.** Equation (2.10) with $f = z_i$, read in two ways: the rank of the five reading lines of $\mathbf{J}$ in equation (11.14), and $\sqrt{\Sigma_{ii}}$ against the prior's standard deviation. Since the coils' two lines are the ones the discharges and the supply share, the coils are also read in the rotated coordinates of mean derate and difference.

**Building blocks.** The Jacobian (11.14) and $\mathbf{\Sigma}$.

**The plant's numbers.** The live reading rows have rank 5 of 7 columns in the closed graph and 4 of 6 in the open, and in both the null columns are exactly $\beta_2$ and $\beta_3$. The open graph has a redundant row: on $(Q, g_1, g_2)$ the four rows $y_{T_r}, y_{T_{d1}}, y_{T_{d2}}, y_Q$ have rank 3, and the row to spare is the heat balance. The closed graph has none: on $(Q, g_1, g_2, b_r)$ the rows $y_{T_s}, y_{T_{d1}}, y_{T_{d2}}, y_Q$ have rank 4 of 4, because $b_r$ has spent the row, and its prior gives it back.

The ratio of posterior to prior spread is 1.000 for $\beta_2$ and $\beta_3$, 0.393 for $\beta_1$, read by $y_{T_1}$, and 0.645 for $b_r$, read through the discharges and the meter. In the rotated coordinates the mean of the coils has 0.0135 against a prior of 0.1061, a ratio of 0.127, and the difference $g_1 - g_2$ has 0.0410 against 0.2121, a ratio of 0.194. By this reading the coils are the best-resolved entries of the hall.

### 11.3.2 What the controller needs

**Control engineer's question.** What does the controller need of them?

**General form.** The moments of a derived quantity, equation (2.10), with the variance (11.1).

**Specialized form.** Equation (2.10) with three functions $f$: the inlet now, $T_3$ of equation (11.6); the inlet at a hold at $T_s = 14°\mathrm{C}$, 200 kW, equation (11.8) with $T_s$ pinned; and the supply a closed node asks for, $T_s^*$ of equation (11.9) at 200 kW, which is what stands against a stop. Differentiating equation (11.6) entry by entry gives the closed gradient, and the open graph's $T_3 = T_s + (Q + Q_o)/\sum G - (1 - \beta_3)(Q + Q_o)/\sum C$ gives the open one, with no column for $b_r$:

$$\mathbf{g}_{T_3}^{\text{closed}} = \begin{bmatrix} -(1-\beta_3)/\sum C \\ 0 \\ 0 \\ 0 \\ 0 \\ (Q+Q_o)/\sum C \\ -1 \end{bmatrix}, \qquad \mathbf{g}_{T_3}^{\text{open}} = \begin{bmatrix} 1/\sum G - (1-\beta_3)/\sum C \\ -(Q+Q_o)\,G_1/(\sum G)^2 \\ -(Q+Q_o)\,G_2/(\sum G)^2 \\ 0 \\ 0 \\ (Q+Q_o)/\sum C \end{bmatrix}. \tag{11.16}$$

**The coils have left the gradient.** Holding the return fixes $T_r$, and the air balance fixes the drop to the cold aisle without any coil. Weak coils are repaid by a colder supply, which the inlet never sees.

**Building blocks.** $\hat{\mathbf{z}}$, $\mathbf{\Sigma}$, the gradients (11.16), and the entry terms of equation (11.1).

**The plant's numbers.** At the MAP state,

$$\mathbf{g}_{T_3}^{\text{closed}} = (-0.0250,\ 0,\ 0,\ 0,\ 0,\ 10.1314,\ -1)^{\mathsf{T}},$$
$$\mathbf{g}_{T_3}^{\text{open}} = (0.0542,\ -8.2679,\ -7.7819,\ 0,\ 0,\ 10.1314)^{\mathsf{T}}.$$

The two gradients are different objects on different states and give the same band, $T_3 = 24.9841°\mathrm{C}$ with sd 0.8283 K, as Section 11.2.3 requires. The band apart from $\beta_3$ is 0.1709 K in both.

Table 11.4 sets the two objects side by side. Each row is an entry or a rotated pair; the second column gives posterior against prior spread, the third the reading that acts on it, the next three the gradient times the spread for each need, and the last the class. Figure 11.2 draws the same numbers for the first two needs: the horizontal axis is the ratio of spreads, so an entry the readings do not reach stands at the right, and the vertical axis is the size of the entry's term, so an entry the quantity needs stands high. Filled marks are the inlet now and open marks the hold; the dashed lines are a guide to the eye, not a test, and they divide the plane into the four classes.

Table 11.4: The entries by what the readings hold on them and by what three needs ask of them. The load has no prior, so its ratio does not exist; it is pinned at both nodes. Terms are $\mathbf{g}_i\sqrt{\Sigma_{ii}}$, in K. The node's error $e$ also enters $T_s^*$, at 0.2598 K on the air balance, and belongs to no entry.

| Entry | Post/prior [–] | Reading | $T_3$ now [K] | Hold, 200 kW [K] | $T_s^*$, 200 kW [K] | Class |
|---|---|---|---|---|---|---|
| $\beta_2$ | 1.000 | none | 0 | 0 | 0 | unobservable, irrelevant |
| $\beta_1$ | 0.393 | $y_{T_1}$ | 0 | 0 | 0 | observable, irrelevant |
| $g_1 - g_2$ | 0.194 | discharges | 0 | −0.0102 | 0.0102 | observable, irrelevant here; a parent's loss |
| mean $g$ | 0.127 | $y_{T_s}$ | 0 | −0.2221 | 0.2221 | observable; the question decides |
| $\beta_3$ | 1.000 | none | 0.8105 | 0.8320 | 0 | **unobservable, needed** |
| $b_r$ | 0.645 | discharges, meter | −0.1935 | 0 | 0 | weakly seen, needed |
| $Q$ | – | $y_Q$ | −0.0692 | 0 | 0 | resolved, slight |

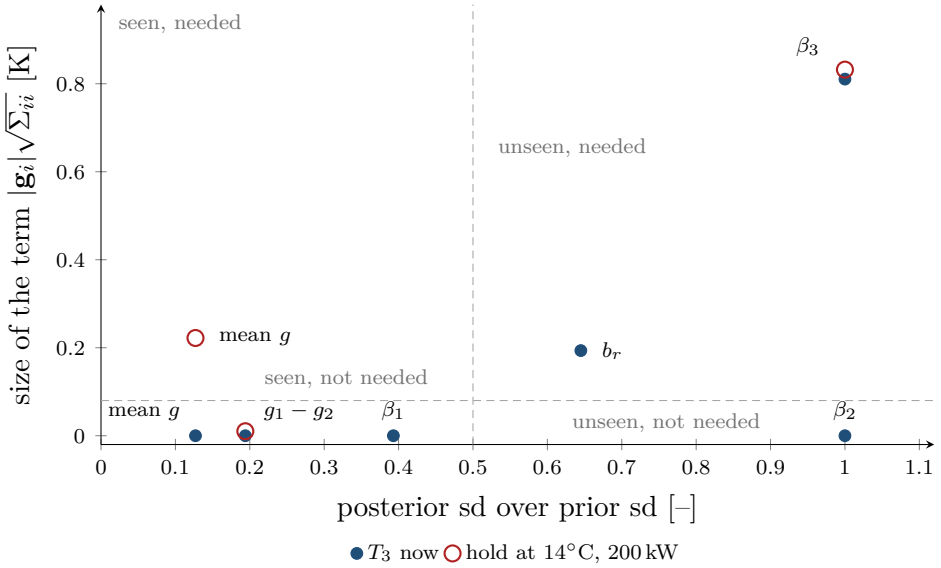



Figure 11.2: The entries of Table 11.4 placed by what the readings hold on them (horizontal) and by what two needs ask of them (vertical). Coordinates are that table's columns, as sizes. Dashed lines are a guide only.

The figure's upper-right corner is the result. The entry the band needs most is one no reading touches, and the best-resolved entries of the hall stand on the floor of the first need. The band's budget, by equation (4.54), in $\mathrm{K}^2$: $\beta_3$ gives 0.6569, $b_r$ gives 0.0374, $Q$ gives 0.0048 and the cross term −0.0130, a total of 0.6861. The class changes with the question: the mean of the coils is irrelevant to the inlet now and supplies nearly all of the state's part of $T_s^*$'s spread,

0.2217 K. With the node's error in equation (11.9) that spread is 0.3416 K on the air balance and 0.3731 K on the reading, so on the air balance the coils give 42 % of its variance. A stop needs the coils and the loop's error together.

Reading only the entry terms misleads for these derived quantities. Against $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$, the root sum of the entry terms is 0.8362 for 0.8283 on the inlet now, 0.8797 for 0.8610 on the hold, and 0.2857 for 0.2217 on the state's part of the stop. The entry view overstates that part by 29 % in this example, because the fit's negative correlation between the coils cancels part of it.

### 11.3.3 The loop with its reading dropped

**Control engineer's question.** The loop's reading drops: what is the loop?

**General form.** Moments at a scenario node, equation (2.10).

**Specialized form.** Equation (2.10) with $f$ the inlet and the return at an open node, $T_s$ frozen. The graph changes as in Chapter 5: the controller's row goes, the supply is pinned, and $b_r$ has no more work to do. Two states of knowledge are possible, and each has its route. The *last belief* keeps the closed fit and reads equation (11.6) at the window's load. The *next window's rows* take the open graph with $T_s$ pinned at 14°C and $y_{T_r}$ removed, and fit again.

**Building blocks.** $\hat{\mathbf{z}}$, $\boldsymbol{\Sigma}$; the open fit with one row fewer; and for a stuck reading the hold at a stop, equation (11.8) with $T_s$ pinned.

**The plant's numbers.** By the last belief, $T_r = 30.0498$°C (sd 0.1935 K) and $T_3 = 24.9841$°C (sd 0.8283 K). By the next window's rows, on the same readings, $T_r = 29.8504$°C (sd 0.2530 K) and $T_3 = 24.8194$°C (sd 0.8338 K). So the frozen loop does not lose the return: the rows left still hold it, at a spread narrower than the sensor's own 0.30 K.

A reading that stays connected and sticks behaves differently, and the loop's error is the reason. Only if the stuck value equals $r$ is the error zero and the integrator stopped, so that the plant is the frozen graph with nothing in the loop's error to show it. That is a knife edge: under fresh errors the last raw value differs from $r$ almost surely, so the generic case is that the integrator walks to a stop. Stuck below $r$ it drives the supply to 20°C, and at 200 kW the inlet is 31.27°C (sd 0.86 K; stated truth 31.69°C); stuck above $r$ it drives the supply to 8°C and the inlet to 19.27°C (truth 19.69°C).

### 11.3.4 The virtual reading

**Control engineer's question.** Can a virtual reading replace it?

**General form.** Moments at a scenario node, equation (2.10), with the reading replaced.

**Specialized form.** Equation (2.10) with $f$ the inlet of the closed node (11.9) and $\sigma_e$ the standard deviation of the route the row is closed on. Each route to $T_r$ is a formula on readings. The air balance is equation (11.7); its spread is $\sqrt{\sigma_{T_d}^2/2 + (\sigma_Q/\sum C)^2}$, with no coil and no fraction in it. The inlet route is $y_{T_1} + (1 - \text{survey } \beta_1)(y_Q + Q_o)/\sum C$, and the fraction's prior enters its spread. The coil route is $y_{T_s} + (y_Q + Q_o)/\text{rated } \sum G$, and the coils' prior on the sum enters its spread.

**Building blocks.** The readings, the tolerances of step 2, $\boldsymbol{\Sigma}$ through the gradient of equation (11.9), which at 200 kW is $(0,\ 0,\ 0,\ 0,\ 0,\ 10.4000,\ 0)^{\mathsf{T}}$: closed at a node, with the load pinned and the error fresh, the state enters the inlet through $\beta_3$ alone.

**The plant's numbers.** Table 11.5 lists the routes, each row one route with its value and spread. Figure 11.3 draws the spreads: each bar's length is the route's standard deviation in K.

Table 11.5: The routes to the return at the window, the first the loop's own.

| Route | Value [°C] | sd [K] |
|---|---|---|
| The reading itself, the loop's | 30.33 | 0.3000 |
| Air balance, mean $y_{T_d} + (y_Q + Q_o)/\sum C$ | 29.8420 | 0.2598 |
| Inlet, $y_{T_1} + (1 - \text{survey } \beta_1)(y_Q + Q_o)/\sum C$ | 30.2216 | 0.8666 |
| Coil, $y_{T_s} + (y_Q + Q_o)/\text{rated } \sum G$ | 28.3600 | 1.5389 |
| The fit, $y_{T_r}$ gone, all rows | 29.8504 | 0.2530 |

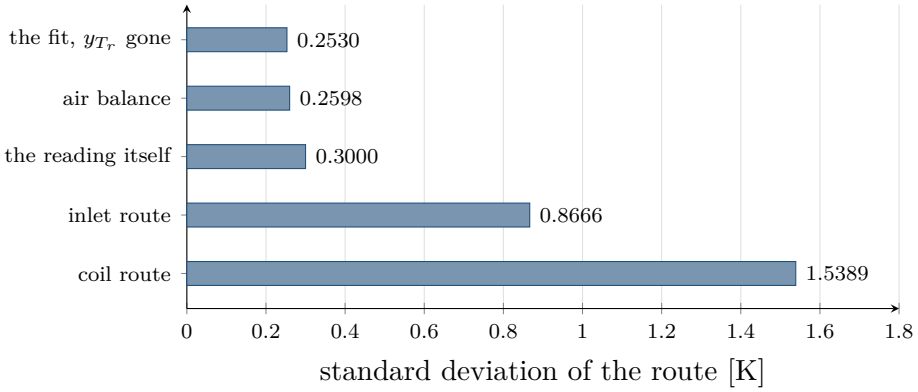



Figure 11.3: The routes of Table 11.5 by standard deviation. The spread of the fit is the fit's with $y_{T_r}$ removed; each other bar is a formula on readings.

The air balance is narrower than the sensor it would replace, and it needs no fit, so it can be computed at the controller. A precision-weighted mean of the two is 0.1964 K and the closed fit's own spread on $T_r$ is 0.1935 K, because that is nearly what the fit does. Three routes taken as independent would give 0.2457 K, but they share $y_Q$, so that is a lower figure. The sibling volume finds the same narrower virtual return on its own plant, which is a statement of direction only.

What the air balance costs is where its errors go. A loop on it holds $T_r = r - e$, and a steady offset on a sensor it reads now passes into the plant: an offset $d$ on one discharge moves $T_r$ by $d/2$, and on the meter by $d/\sum C$. At 0.30 K on one discharge, $T_r$ and $T_3$ move by 0.150 K; at 3 kW on the meter, by 0.150 K. Closing on the route spends one combination of those three rows, as Chapter 5's loop spent its own reading: the combination's error becomes the loop's error, and the other two combinations stay live, for example $y_{T_{d1}} - y_{T_{d2}}$. The band counts nothing twice, because $T_3^*$'s gradient is $\beta_3$ alone and 0.2598 K is the error's only entry. A fit that kept all the rows and added a prior on $e$ would count them twice, taking 0.2530 to 0.1812 K.

At 200 kW, with zero-mean error at the node, the loop on the air balance has $T_3^* = 25.130$°C, sd 0.8716 K, and the loop on the physical reading has the same mean, sd 0.8844 K. The small gain is what the chapter's $\beta_3$ leaves room for. The choice of a fresh error was tested, and the test moves the mean as well as the band. If the error persists at the fit's value, its spread at the node is the posterior's 0.1935 K for either loop, and sd $T_3^*$ is 0.8542 K. The reading's loop then carries $e = \hat{b}_r = +0.2802$ K, with $T_3^* = 24.850$°C, $P = 0.9941$. The air balance carries $e = v - \hat{T}_r = -0.2078$ K, with $T_3^* = 25.338$°C, $P = 0.9742$. Under that test the air balance goes from slightly better to 0.488 K warmer, the move of the next sentence. At the switch, with the set point kept, the reading and the air balance differ by 0.4880 K, which is 1.23 standard deviations of their difference (0.3969 K), so the supply moves by that at steady state. How the

move is made is the bumpless transfer of Chapter 27.

### 11.3.5 The hold against the virtual loop as the load moves

**Control engineer's question.** Hold or virtual loop, as the load moves?

**General form.** The chance constraint (6.1) of Chapter 6, the threshold form (2.12), with its normal reading (6.3), read at each load, and inverted for the load and for the setting.

**Specialized form.** Equation (6.1), itself equation (2.12), with $f$ the inlet of the frozen node (11.8) or the closed node (11.9), $f^* = 27°\mathrm{C}$ and $\alpha = 0.05$. Before any spread, what separates the two nodes is the sign of the inlet's slope in the pinned load. From the two equations,

$$\left.\frac{\partial T_3^*}{\partial Q^*}\right|_{\text{closed}} = -\frac{1-\beta_3}{\sum C}, \qquad \left.\frac{\partial T_3^*}{\partial Q^*}\right|_{\text{frozen}} = \frac{1}{\sum G} - \frac{1-\beta_3}{\sum C}. \tag{11.17}$$

The term $-(1-\beta_3)/\sum C$ is in both, and it is the air balance: more load opens the drop between return and cold aisle, and rack 3, which breathes only the fraction $\beta_3$ of the return, stands $(1-\beta_3)$ of that drop below the return. The frozen node adds $1/\sum G$, the coils' rise, since the return floats up when the supply does not move. The closed node has no such term, because the controller's row pins the return and pays for the load out of the supply. A return-air loop therefore cools the rack as load rises, with the airflow pinned. Return air is held because it is one mixed stream, which a single sensor in one duct can read for the whole hall, where the rack inlets are many and scattered.

The signs do not depend on this plant's $\beta_3$ or coils. For every $\beta_3$ in $[0, 1)$ the closed slope is negative and the frozen slope positive, 0.0292 to 0.0787 K/kW at the fit's $\sum G$ for $\beta_3$ from 0 to 0.99. The closed slope has no coil in it, so coils cannot alter it, and the frozen slope flips only if $\sum G > \sum C/(1-\beta_3)$, which at $\beta_3 = 0.5$ is 40 against 12.63 kW/K. The condition doing the work is pinned airflow: if airflow tracks load, the drop $(Q + Q_o)/\sum C$ is constant and the closed slope is 0.

The band is a function of the load as well. In the closed node the gradient entry on $\beta_3$ is $(Q^* + Q_o)/\sum C$; with the fresh error, $\mathrm{Var}(T_3^* \mid \mathbf{y}) = ((Q^* + Q_o)/\sum C)^2\, \Sigma_{\beta_3\beta_3} + \sigma_e^2$. The chance constraint for the set point at one load inverts directly, since $r$ is additive; for a hold, $T_s$ is additive in the same way, as in equation (4.66).

**Building blocks.** $\hat{\mathbf{z}}$, $\boldsymbol{\Sigma}$, the node's gradient at each load, $\Phi^{-1}(0.95) = 1.6449$, and the stops.

**The plant's numbers.** At the fit's mean the two slopes are $-0.0250$ K/kW closed, and $0.0792 - 0.0250 = 0.0542$ K/kW frozen. So the two break on opposite sides. Table 11.6 gives the two nodes at $r = 30.33°\mathrm{C}$. Each row is a load; the columns give the frozen node's mean with its spread and probability, the air-balance loop's mean with its spread and probability, the supply that loop asks for, and the stated truth for each. A dagger marks a closed node past a stop, whose values describe no plant. Figure 11.4 draws the same results. The horizontal axis is the load and the vertical axis the inlet; the solid lines are the means, the shaded bands run from the mean to the mean plus 1.6449 standard deviations, so a band's upper edge above the red limit line breaks the constraint; the gray regions are where the closed node is void, so the loop's lines stop there; the vertical dotted lines mark the two crossings; and the black dotted lines are the stated truth, the closed one drawn only inside its own stops.

- *The crossing.* Above 205.10 kW the frozen actuator breaks 0.95, and the virtual loop holds: at that load it has $P = 0.9875$ with $T_s^* = 13.45°\mathrm{C}$, valid, up to $T_s$ at 8°C, at 273.94 kW. The truth's frozen crossing is 223.24 kW.

Table 11.6: The inlet at each load, $r = 30.33°\text{C}$, by the state route. $P$ is $\Pr(T_3^* \le 27°\text{C} \mid \mathbf{y})$, the loop's at zero-mean error at the node. †: past a stop, void.

| $Q^*$ [kW] | Frozen $T_3^*$ [°C] | $P$ [–] | Loop $T_3^*$ [°C] | $P$ [–] | Loop's $T_s^*$ [°C] | Truth frozen [°C] | Truth loop [°C] |
|---|---|---|---|---|---|---|---|
| 120 | 20.94 (0.53) | 1.000 | 27.13 (0.57)† | void | 20.19† | 21.20 | 27.58† |
| 140 | 22.02 (0.61) | 1.000 | 26.63 (0.65) | 0.716 | 18.61 | 22.32 | 27.15 |
| 160 | 23.11 (0.69) | 1.000 | 26.13 (0.72) | 0.886 | 17.02 | 23.45 | 26.72 |
| 180 | 24.19 (0.78) | 1.000 | 25.63 (0.80) | 0.958 | 15.44 | 24.57 | 26.29 |
| 200 | 25.27 (0.86) | 0.978 | 25.13 (0.87) | 0.984 | 13.86 | 25.69 | 25.86 |
| 220 | 26.36 (0.94) | 0.752 | 24.63 (0.95) | 0.994 | 12.27 | 26.82 | 25.43 |
| 240 | 27.44 (1.03) | 0.333 | 24.13 (1.03) | 0.997 | 10.69 | 27.94 | 25.00 |
| 260 | 28.53 (1.11) | 0.084 | 23.63 (1.10) | 0.999 | 9.10 | 29.07 | 24.57 |
| 280 | 29.61 (1.19) | 0.014 | 23.13 (1.18)† | void | 7.52† | 30.19 | 24.14† |

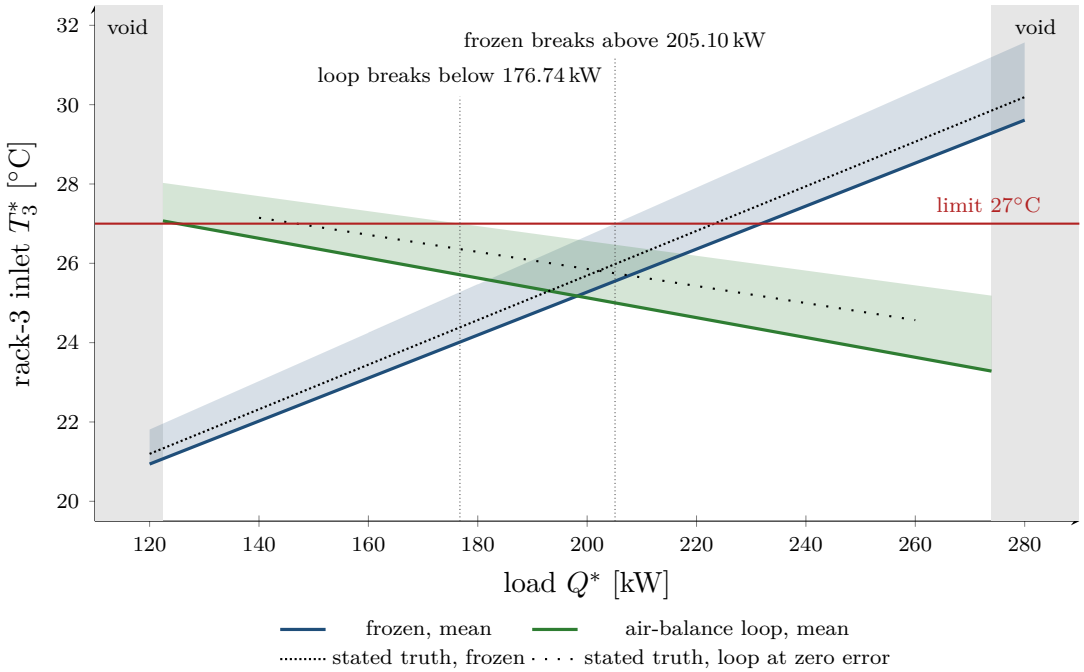



Figure 11.4: The inlet against the load at $r = 30.33°\text{C}$, frozen (equation (11.8)) and closed on the air balance (equation (11.9)), means and bands up to the 0.95 quantile from Table 11.6. The closed node is void below 122.43 kW and above 273.94 kW. The loop's crossing assumes zero-mean error at the node; with the fit's error persisting it is 186.18 kW.

- *The reverse crossing.* Below 176.74 kW, a figure that assumes zero-mean error at the node, the virtual loop breaks 0.95, with $T_s^* = 15.70$°C, valid, and the frozen actuator holds with $P = 1.0000$. The loop on the physical reading crosses at 177.97 kW, and the truth's loop, at zero error, at 146.88 kW, a valid node at $T_s^* = 18.29$°C.

- *With the fit's error persisting.* The air balance crosses at 186.18 kW ($T_s^* = 15.16$°C) and the physical reading at 160.00 kW ($T_s^* = 16.74$°C), so the air balance is 26 kW worse, where at zero-mean error it was about 1 kW better.

- *What the crossing is made of.* Mostly the survey's band. At the mean the loop reaches 27°C at 125.20 kW; with $e = 0$, $\beta_3$'s band alone moves that to 172.78 kW, and the error's band to 176.74 kW. The breach does not extend downward without end: $T_s$ reaches 20°C at 122.43 kW, and pinned there the plant regains 0.95 below 106.73 kW.

**The loop on the reading does the same before any loss.** Its probability at 180 kW is 0.9547 and at 160 kW is 0.8814. At zero-mean error the virtual loop inherits this, and the loss did not cause it. The engineer's margin of 3.34 K was a statement about one load on one sheet.

Table 11.7 inverts the constraint for the setting. Each row is a load; the columns give the highest return set point on the physical reading, and on the air balance with the supply it asks for, and the highest supply of a hold. All these nodes are valid. Over the stated range of 160 to 240 kW one setting has to hold at the worst load for it: the air-balance loop's $r = 30.015$°C binds at 160 kW, and it runs the supply at 16.71, 13.54 and 10.37°C at 160, 200 and 240 kW; the hold's $T_s = 11.869$°C binds at 240 kW. Neither dominates. The inlet means of the loop are 25.815, 24.815 and 23.815°C, and of the hold 20.975, 23.143 and 25.311°C: the hold runs warmer at 240 kW, the loop at 160 and 200 kW. The hold's $P$ is above 0.9999 at 160 and at 200 kW and 0.950 at 240 kW, and it spends no sensors. The loop's setting assumes zero-mean error at the node; with the fit's error persisting, the air balance's setting at 160 kW is 29.842°C.

Table 11.7: The settings that hold 0.95 at each load, by equation (6.3), at zero-mean error at the node, with the stated truth's largest values at zero error.

| $Q^*$ [kW] | $r_{95}$, reading [°C] | $r_{95}$, air balance [°C] | its $T_s^*$ [°C] | $T_{s,95}$, hold [°C] | Truth $r$ [°C] | Truth $T_s$ [°C] |
|---|---|---|---|---|---|---|
| 160 | 29.990 | 30.015 | 16.709 | 16.750 | 30.612 | 17.555 |
| 180 | 30.368 | 30.391 | 15.501 | 15.530 | 31.042 | 16.431 |
| 200 | 30.745 | 30.766 | 14.292 | 14.310 | 31.472 | 15.307 |
| 220 | 31.121 | 31.140 | 13.082 | 13.090 | 31.902 | 14.182 |
| 240 | 31.495 | 31.513 | 11.871 | 11.869 | 32.332 | 13.058 |

One comparison lies outside the questions and is given as a pointer. A loop on a virtual *inlet*, mean $y_{T_d} + \hat{\beta}_3(y_Q + Q_o)/\sum C$ with reading sd 0.2250 K, has one setting over the range at $r_3 = 25.327$°C, and the supply runs 16.22, 14.05 and 11.89°C. It stands near the hold at 240 kW and near the return loop at 160 kW, since it reads the protected quantity. Choosing the set point is Chapter 16's.

### 11.3.6 The loss of each other reading

**Control engineer's question.** Which other reading matters most?

**General form.** Moments with a row removed, equation (2.10).

**Specialized form.** Equation (2.10) with $f = T_3$ or the node's $T_3^*$, under the posterior without the row, reached by one of two routes that must not be mixed. The *down-date* keeps $\hat{\mathbf{z}}$ and

removes the row's information, $\boldsymbol{\Sigma}_{-k} = (\boldsymbol{\Sigma}^{-1} - \mathbf{J}_k^{\mathsf{T}}\mathbf{J}_k)^{-1}$. The *refit* removes the row and solves again, so the state moves too. With a virtual loop already running, the third route is a formula: which route to $T_r$ remains, and with what spread, in equation (11.9).

**Building blocks.** $\boldsymbol{\Sigma}$, the lines of $\mathbf{J}$, the node's gradient. For the air-balance loop after a second loss: if the meter goes, both balances give $T_r = (\sum C\,T_c - \sum G\,y_{T_s})/(\sum C - \sum G)$ with $\sum G$ at its last belief; if a discharge goes, the other gives $T_r = y_{T_d} + s_u(y_Q + Q_o)/C_u$, with $s_u$ the unit's share of the heat at its last belief, 0.4849 (sd 0.0103) for unit 2.

**The plant's numbers.** At the window, the two graphs agree to the digits printed, by the identity. Table 11.8 lists the added band on $T_3$. Each row is a removed reading; the columns give the added band by refit, the added band by down-date, and what is left apart from $\beta_3$ by refit, against the present 0.1709 K. The routes rank the meter differently, and the reason is not information. The refit moves $\hat{Q}$ to 203.73 kW, and $\beta_3$'s term in the band grows with the state; the state moved, and no information did. So a ranking names its route.

Table 11.8: Added band on $T_3$ at the window, from sd 0.8283 K, by each route. $y_{T_r}$ exists only in the open graph and $y_{T_s}$ only in the closed.

| Reading removed | Refit [K] | Down-date [K] | Apart from $\beta_3$, refit [K] |
|---|---|---|---|
| $y_Q$ | +0.0377 | +0.0021 | 0.1811 |
| $y_{T_{d2}}$ | +0.0239 | +0.0273 | 0.2734 |
| $y_{T_{d1}}$ | +0.0233 | +0.0276 | 0.2744 |
| $y_{T_r}$, open only | +0.0055 | +0.0109 | 0.2175 |
| $y_{T_1}$ | +0.0007 | +0.0007 | 0.1741 |
| $y_{T_s}$, closed only | −0.0007 | +0.0001 | 0.1715 |

What the loop changes appears only at a node. Figure 11.5 draws it. The horizontal axis lists the readings, the vertical axis is the added band, and the four series are four situations: at the window by down-date; at the node of a loop on the physical reading, its error persisting, by down-date; at the node of the air-balance loop, 200 kW, with $y_{T_r}$ already gone, by the route that remains; and at the frozen node, 240 kW, by down-date. Within one series the route is one; between series only the direction is read.

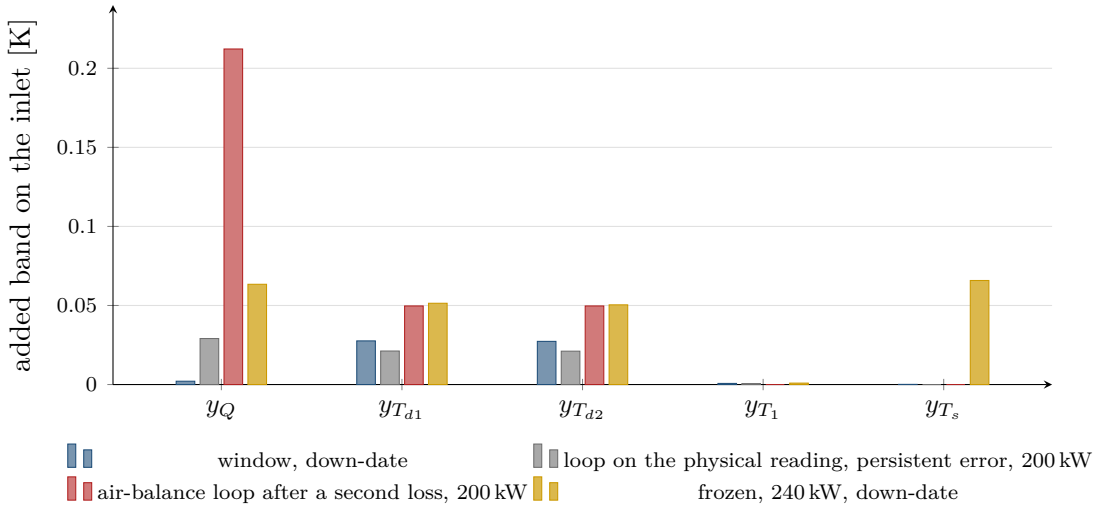



Figure 11.5: The band each lost reading adds, in four situations. The window's values are Table 11.8's second column.

On the physical reading's loop with a fresh error, removing any other reading adds 0.0000 K to 0.8844 K, but that is the fresh draw's doing: the node then holds no entry but $\beta_3$. With the error persisting the node reads $b_r$, and from 0.8542 K the meter adds 0.0291, the discharges 0.0212 and 0.0211, $y_{T_1}$ 0.0006 and $y_{T_s}$ 0.0000 K. On the air-balance loop the rows that were harmless become loop-dependent. Each second loss is absorbed by another route, at a price. Losing the meter adds 0.2122 K, by the route of both balances at 0.6946 K, sd $T_3^*$ 1.0839 K. Losing either discharge adds 0.0497 K, by one discharge and the split at 0.3958 K, sd $T_3^*$ 0.9213 K. At the frozen node at 240 kW, where the coils now matter, $y_{T_s}$ adds 0.0658 K, $y_Q$ 0.0634, $y_{T_{d1}}$ 0.0514, $y_{T_{d2}}$ 0.0504 and $y_{T_1}$ 0.0009 K. The supply sensor, worth nothing at the window, ranks first, because the frozen node's gradient, $(-0.0792,\ -1.8512,\ -1.7424,\ 0,\ 0,\ 12.4000,\ -1)^{\mathsf{T}}$, reads the state through the window's supply.

The sibling volume's Chapter 17, on its two-hall plant, finds that the return can go first and that a discharge costs more after it. The direction holds here, by the like computation of the open graph with $T_r$ as the quantity and the down-date as the route: losing the return adds 0.0597 K, against 0.1032 for the meter and 0.0787 for $y_{T_{d1}}$, and after the return is out $y_{T_{d1}}$ adds 0.3946 K. On $T_3$ the same discharge goes from +0.0276 to +0.1908 K. The plants differ, so no number compares.

### 11.3.7 Belief aggregation

**Control engineer's question.** What does a parent level keep?

**General form.** The moments of a derived quantity, equation (2.10), here a vector $\mathbf{a}(\mathbf{z})$.

**Specialized form.** Equation (2.10) with the aggregate map $\mathbf{a}(\mathbf{z}) = (Q + Q_o,\ G_1 + G_2)^{\mathsf{T}}$, the hall's heat and its total conductance. Taking coordinates $\mathbf{w} = (\mathbf{a},\ g_1 - g_2,\ \beta_1, \beta_2, \beta_3, b_r)$ and writing the information in them as blocks, the marginal covariance is the inverse of the Schur complement, as Chapter 3 derived it,

$$\boldsymbol{\Sigma}_{\mathbf{a}} = \left(\boldsymbol{\Lambda}_{\mathbf{aa}} - \boldsymbol{\Lambda}_{\mathbf{ab}}\boldsymbol{\Lambda}_{\mathbf{bb}}^{-1}\boldsymbol{\Lambda}_{\mathbf{ba}}\right)^{-1}, \tag{11.18}$$

which equals the delta route with the map's $2 \times 7$ Jacobian $\mathbf{A}$, $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathsf{T}}$.

**Building blocks.** $\boldsymbol{\Sigma}$, $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}^{-1}$ and the map's Jacobian.

**The plant's numbers.** Table 11.9 lists what the map keeps and what it loses, each row a quantity with the unit posterior, the parent's value and the truth. Figure 11.6 draws the loss. Its axes are the coils' log-derates; the ellipses are the one- and two-sd contours of the two-unit posterior; the three curves are constant $G_{\text{tot}}$ at the mean and one sd either side, which is all the parent holds on the coils; the long bar is the parent's equal split with $g_1 - g_2$ back at its prior, one prior sd each way; the dotted diagonal separates which unit is more derated; and the cross is the stated truth.

The covariance by the Schur complement (11.18) equals the delta route to $7 \times 10^{-15}$, which is one piece of linear algebra done twice; the draws of Section 11.4.2 test the map's curvature. The parent keeps the hall's totals and little else. Along the curve of constant $G_{\text{tot}}$ it has lost which unit is derated, and that decides what a trip leaves: $G_1$ is 6.504 against the parent's 6.313 kW/K, with a spread more than four times wider. At this load the tripped node, on $T_r = r$, is void, with $T_s^* = -0.82$°C past $T_s$ at 8°C, so the useful figure is the load at which $G_1$ holds $r$ at 8°C. The inlet stays unknown to the parent as to the hall: the aggregate takes its spread only from 0.8283 to 0.8119 K. Chapter 30 takes the map up, as the description a hall passes to a site level whose decisions run slower than the loop; this section fixes only what that description keeps.

Table 11.9: The aggregate **a** and what it no longer holds, with standard deviations in parentheses.

| Quantity [unit] | Two-unit posterior | Parent | Truth |
|---|---|---|---|
| $Q_{\mathrm{tot}}$ [kW] | 202.63 (2.768) | the same | 208.00 |
| $G_{\mathrm{tot}}$ [kW/K] | 12.626 (0.170) | the same | 12.867 |
| Correlation of the two [–] | 0.579 | the same | – |
| $g_1 - g_2$ [–] | 0.0606 (0.0410) | prior, sd 0.2121 | 0.0700 |
| Pr(unit 2 the more derated) [–] | 0.930 | 0.5 | – |
| $G_1$, left if unit 2 trips [kW/K] | 6.504 (0.156) | equal split 6.313 (0.675) | 6.659 |
| Load where $G_1$ holds $r$ at 8°C [kW] | 137.24 | 132.97 | 140.69 |
| sd $T_3$ [K] | 0.8283 | given **a**, 0.8119; given return only, 0.8128 | – |

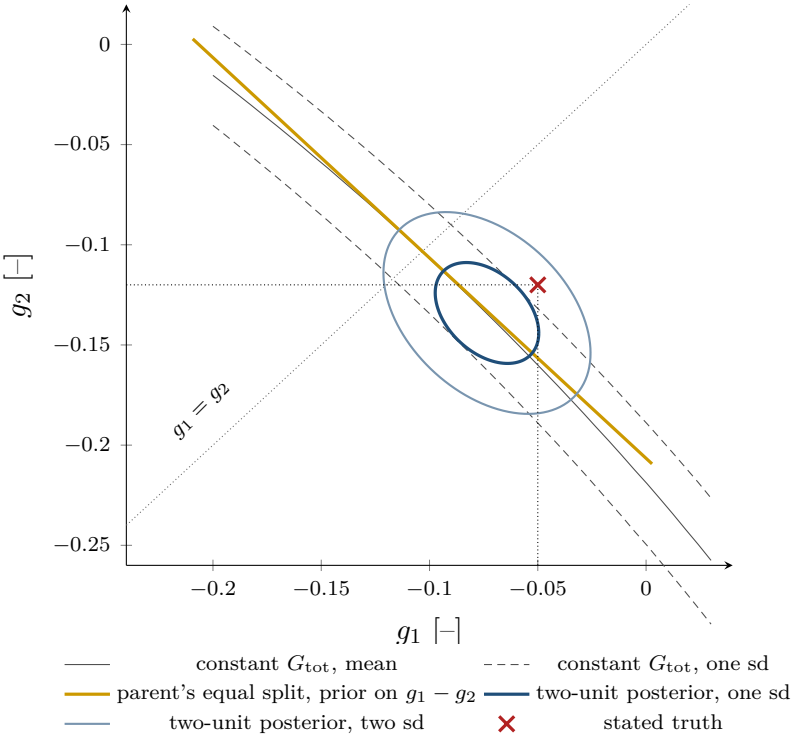



Figure 11.6: The coils at the two-unit posterior against what the aggregate keeps. Contours: the Laplace normal with spreads 0.02391 and 0.02518 and correlation −0.3977. Curves: $G_{\mathrm{tot}} = 7(\mathrm{e}^{g_1} + \mathrm{e}^{g_2})$ at 12.6261 kW/K and ±0.1700.

## 11.4 The checks (step 5)

The checks test the answers against the full state, the draws, the record's stated truth and the stated priors. The stated tolerances are 0.005 on a probability, 0.5 kW on a crossing and $10^{-3}$ on the identity.

### 11.4.1 The limit of exact physical rows

Table 11.10 solves the full state, with the seven temperatures of the chain added and the physical and controller's rows at tolerance $\sigma_p$; $T_2$ is read by nothing and is left out. Each row is a tolerance, the last the reduced state. By $\sigma_p = 0.001$ the two agree to the printed digits.

Table 11.10: The full state at three tolerances against the reduced state.

| $\sigma_p$ [K] | $T_3$ (sd) [°C (K)] | $b_r$ (sd) [K] | $g_1$ (sd) [–] |
|---|---|---|---|
| 0.100 | 24.9740 (0.8347) | 0.2631 (0.2016) | −0.0731 (0.0240) |
| 0.010 | 24.9840 (0.8284) | 0.2800 (0.1935) | −0.0735 (0.0239) |
| 0.001 | 24.9841 (0.8283) | 0.2802 (0.1935) | −0.0735 (0.0239) |
| Reduced | 24.9841 (0.8283) | 0.2802 (0.1935) | −0.0735 (0.0239) |

### 11.4.2 The state route against reweighted draws

Table 11.11 compares the state route with reweighted draws, the proposal a $t_5$ on the Laplace normal with seed 11 and $N = 10^6$, effective size 779 835. Each row is a probability; the columns give the two values, the Monte Carlo standard error and the gap. The largest gap is −0.0015, at the frozen node at 240 kW. It is three standard errors, so it is not the draws; it is the map's curvature, inside the tolerance. The frozen crossing by the draws is 204.98 kW against 205.10 kW. The aggregate by the draws has spreads 2.771 kW and 0.170 kW/K and correlation 0.579, against the delta route's 2.768, 0.170 and 0.579.

Table 11.11: $\Pr(T_3^* \le 27°\text{C} \mid \mathbf{y})$ by the state route and by reweighted draws.

| Node | State route [–] | Draws [–] | se [–] | Gap [–] |
|---|---|---|---|---|
| Frozen, 205.10 kW | 0.9500 | 0.9492 | 0.0002 | −0.0008 |
| Air-balance loop, 176.74 kW | 0.9500 | 0.9500 | 0.0002 | +0.0000 |
| Frozen, 240 kW | 0.3327 | 0.3311 | 0.0005 | −0.0015 |
| Air-balance loop, 240 kW | 0.9974 | 0.9974 | 0.0000 | +0.0000 |
| Frozen, 200 kW | 0.9776 | 0.9774 | 0.0002 | −0.0002 |

### 11.4.3 The answers against the stated truth

The record is built from a stated truth, whose offset is $b_r = 0.1647$ K. Table 11.12 compares it with the fit; each row is an entry or a derived temperature, and the last column is the distance in the fit's standard deviations. The largest is the load's, +1.94. The inlet's truth is 25.69°C against the fit's 24.98°C, +0.86 standard deviations, and most of that is $\beta_3$ at +0.88, the entry nothing reads. At the crossings the truth is on the safe side of the limit: at 205.10 kW the frozen plant has a true inlet of 25.9798°C, and at 176.74 kW a zero-error loop has 26.3582°C, with $P = 0.9933$ over the fresh error.

Table 11.12: The fit against the stated truth of the constructed record.

| Entry or temperature [unit] | Fit | sd | Truth | Truth minus fit [sd] |
|---|---|---|---|---|
| $Q$ [kW] | 194.6273 | 2.7676 | 200.0000 | +1.94 |
| $g_1$ [–] | −0.0735 | 0.0239 | −0.0500 | +0.98 |
| $g_2$ [–] | −0.1340 | 0.0252 | −0.1200 | +0.56 |
| $\beta_1$ [–] | 0.2204 | 0.0314 | 0.2200 | −0.01 |
| $\beta_2$ [–] | 0.3500 | 0.0800 | 0.3100 | −0.50 |
| $\beta_3$ [–] | 0.5000 | 0.0800 | 0.5700 | +0.88 |
| $b_r$ [K] | 0.2802 | 0.1935 | 0.1647 | −0.60 |
| $T_s$ [°C] | 14.0014 | 0.0500 | 14.0000 | −0.03 |
| $T_r$ [°C] | 30.0498 | 0.1935 | 30.1653 | +0.60 |
| $T_c$ [°C] | 19.9184 | 0.1748 | 19.7653 | −0.88 |
| $T_3$ [°C] | 24.9841 | 0.8283 | 25.6933 | +0.86 |

### 11.4.4 The priors the band rests on

Table 11.13 varies the two stated priors. Each row is a width; the columns give the mean and spread on $T_3$, for $b_r$ the frozen crossing, and for $\beta_3$ the air-balance loop's crossing. The offset's prior hardly moves the spread, but it moves the means: from 0.1 to 0.6 K, $\hat{b}_r$ runs 0.0647 to 0.4072 K, $\hat{T}_3$ runs 25.16 to 24.88°C and the frozen crossing 203.55 to 206.03 kW. The row identity of Section 11.2.3 holds only at 0.30 K, the reading's tolerance, which is the reason to keep that width. The survey's tolerance is the band. At twice the tolerance the crossing, 275.28 kW, lies above $T_s$ at 8°C, at 273.94 kW, so at that prior the loop holds at no valid node. **The band is a survey's tolerance**, and it has to be stated with the answer.

Table 11.13: The band and the two crossings over the stated priors.

| Prior, sd | $\hat{T}_3$ [°C] | sd $T_3$ [K] | Frozen crossing [kW] | Loop's crossing [kW] |
|---|---|---|---|---|
| $b_r$, 0.10 K | 25.1621 | 0.8223 | 203.55 | – |
| $b_r$, 0.30 K | 24.9841 | 0.8283 | 205.10 | – |
| $b_r$, 0.60 K | 24.8792 | 0.8318 | 206.03 | – |
| $\beta_3$, 0.04 | – | 0.4398 | – | 152.36 |
| $\beta_3$, 0.08 | – | 0.8283 | – | 176.74 |
| $\beta_3$, 0.16 | – | 1.6300 | – | 275.28 |

## 11.5 The report (step 6)

At the window the loop changed nothing in the band: the inlet is 24.98°C with sd 0.83 K in the open and the closed graph, and that band is the survey's $\beta_3$, which no reading holds. What the loop changes is the plant left after the loss. A supply frozen at 14°C breaks 0.95 above 205.10 kW, and a loop closed on the air balance breaks it below 176.74 kW at zero-mean error, or 186.18 kW if the fit's error persists, at the same set point, because with the airflow pinned a return-air loop cools the rack as load rises. A hold reset to 11.87°C holds the range of 160 to 240 kW. The sheet's margin of 3.34 K decides neither.

### 11.5.1 The answers and their bands

Table 11.14 gives the answers in the order of Table 11.1, by the state route unless a row names another. Each row is a question; the columns give the answer, its band and the stated truth.

Table 11.14: The report for the two-unit hall and its loop, $r = 30.33$°C. Draws: $t_5$ proposal, seed 11, $N = 10^6$.

| Control engineer's question | Answer | Band | Stated truth |
|---|---|---|---|
| What can these readings resolve? | rank 5 of 7 closed, 4 of 6 open; $\beta_2$, $\beta_3$ unseen in both | post/prior: mean $g$ 0.127, $g_1 - g_2$ 0.194, $\beta_1$ 0.393, $b_r$ 0.645 | $\beta_3 = 0.57$; $b_r = 0.1647$ K |
| What does the controller need of them? | for $T_3$: $\beta_3$, $b_r$; the coils have left the gradient. For a stop: the mean of the coils and the node's error | $T_3 = 24.9841$°C, sd 0.8283 K; $T_s^*$ sd 0.2217 K from the state, 0.3416 K with $e$ on the air balance | 25.6933°C |
| The loop's reading drops: what is the loop? | next window's rows, frozen: $T_r = 29.8504$°C, $T_3 = 24.8194$°C | sd 0.2530, 0.8338 K | 30.1653, 25.6933°C |
| Can a virtual reading replace it? | air balance 29.8420°C, at 200 kW $T_3^* = 25.130$°C; if the fit's error persists, 25.338°C, and the reading's loop 24.850°C | sd 0.2598 K, narrower than the sensor's 0.30; sd $T_3^*$ 0.8716 K; persistent, 0.8542 K | 25.86°C |
| Hold or virtual loop, as the load moves? | frozen breaks above 205.10 kW (draws 204.98); loop breaks below 176.74 kW at zero-mean error, 186.18 kW persistent (reading's loop 160.00 kW), void past 122.43 and 273.94 kW; a hold at 11.87°C holds 160 to 240 kW | $\beta_3$ prior 0.04 to 0.16: crossing 152.36 to 275.28 kW | 223.24; 146.88 kW |
| Which other reading matters most? | at the window, by refit, $y_Q$ +0.0377 K; after the air-balance loop, $y_Q$ +0.2122 K, a discharge +0.0497 K; frozen at 240 kW, $y_{T_s}$ +0.0658 K; reading's loop with persistent error, $y_Q$ +0.0291 K | routes named in Table 11.8; the down-date gives $y_Q$ +0.0021 K | – |
| What does a parent level keep? | (202.63 kW, 12.626 kW/K), correlation 0.579; loses the split | sd 2.768 kW, 0.170 kW/K | (208.00, 12.867) |
| Which sensor next? | permanent rack-3 inlet (Section 11.6.1) | sd $T_3$ from 0.8283 to 0.0861 K | – |

### 11.5.2 The entry in the commissioning log

On losing the return sensor, manual at the last supply fails above 205 kW; a hold reset to 11.87°C holds the range of 160 to 240 kW, with $P$ above 0.9999, above 0.9999 and 0.950 at 160, 200 and 240 kW, and it spends no sensors. A loop closed on the air balance also serves above 205 kW, with two conditions stated. Its stability has not been computed. And the two discharges and the meter become loop-dependent, the meter most of all. The same set point on the air balance fails below 177 kW, or below 186 kW if its error persists, so the entry records the load range the setting was checked over. A loop on the physical reading has that low-load behavior too, and the entry says so. The inlet figures carry the survey's tolerance on rack 3, and the entry states it. The verdict rests on one window's draw of its sensors' errors. The set point stated here is the window's; choosing a set point is left to Chapter 16, the transfer to Chapter 27.

### 11.5.3 Common mistakes

Table 11.15 lists five mistakes this plant makes easy. Each row names the mistake, its form on this hall and the correction.

Table 11.15: Five mistakes in reasoning about a lost sensor under feedback.

| Mistake | On this hall | Correction |
|---|---|---|
| Ranking sensors by how well each entry is resolved | the reference load meter tops the entry spread at 2.5549 and leaves sd $T_3$ at 0.8280 K | rank by $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ of the quantity the controller needs |
| Going to manual at the last supply because it feels conservative | frozen at 14°C, $P = 0.333$ at 240 kW against the loop's 0.997; reset to 11.87°C, 0.950 | compute both nodes; they break on opposite sides, and a hold is reset for the range |
| Reading a closed node past a stop | $T_3^* = 23.13$°C at 280 kW, with $T_s^* = 7.52$°C | mark the node void; at the stop the actuator is pinned again |
| Comparing a down-date with a refit | $y_Q$: +0.0021 against +0.0377 K, the second with $\hat{Q}$ at 203.73 kW | name the route; a refit moves the state as well as the information |
| Taking the virtual reading as free | narrower at 0.2598 K, yet 0.30 K on one discharge moves the inlet 0.150 K | state the sensors the loop now depends on; another route absorbs each loss, at a wider band |

## 11.6 The next measurement (step 7)

The floor of the band is an entry no reading reaches, so the next measurement has to reach it. A new row returns to step 3, and what it would remove is known before its value is.

### 11.6.1 The sensor ranked by need and by spread

**Control engineer's question.** Which sensor next?

**General form.** The design form (2.16) over candidate rows, through the preposterior (4.73), with a score on the result.

**Specialized form.** Equation (2.16), through equation (4.73), with each candidate's line $\mathbf{J}_a$, scored two ways: by the controller's need, $\sqrt{\mathbf{g}_{T_3}^{\mathsf{T}}\boldsymbol{\Sigma}_a\mathbf{g}_{T_3}}$ with $\mathbf{g}_{T_3}$ of equation (11.16); and by entry spread, $\sum_i(\sqrt{\Sigma_{a,ii}}/\text{prior}_i)^2$. A rack-3 inlet sensor has the line $\mathbf{g}_{T_3}^{\mathsf{T}}/0.3$. The load has no prior, so $Q$ enters the spread score on its own present spread, and that base is why the meter tops that column.

**Building blocks.** **Σ**, the candidates' lines, the need's gradient. The permanent rack-3 sensor averages twelve windows of independent error, 0.087 K.

**The plant's numbers.** At present sd $T_3$ is 0.8283 K and the spread score 3.6240. Table 11.16 lists the ten candidates, each row with its spread on $T_3$, its rank by need, its score by entry spread and its rank by that.

Table 11.16: The candidates ranked by the controller's need and by entry spread. Ties in the score break by name.

| Candidate | sd $T_3$ [K] | Need rank [–] | Spread score [–] | Spread rank [–] |
|---|---|---|---|---|
| Rack-3 inlet, permanent, 12 windows, 0.087 K | 0.0861 | 1 | 2.6590 | 2 |
| Repeat survey of $\beta_3$, 0.02 | 0.2605 | 2 | 2.6828 | 4 |
| Rack-3 inlet, 0.3 K | 0.2821 | 3 | 2.7616 | 5 |
| Reference return, 0.05 K | 0.8137 | 4 | 2.9880 | 7 |
| Second return, 0.3 K, not in the loop | 0.8238 | 5 | 3.4247 | 8 |
| Cold aisle, 0.3 K | 0.8245 | 6 | 3.5411 | 10 |
| Rack-2 inlet, 0.3 K | 0.8277 | 7 | 2.7633 | 6 |
| Reference load meter, 0.5 kW | 0.8280 | 8 | 2.5549 | 1 |
| Unit-1 heat meter, 3 kW | 0.8283 | 9 | 3.4742 | 9 |
| Repeat survey of $\beta_2$, 0.02 | 0.8283 | 10 | 2.6828 | 3 |

The two rankings disagree, and a like comparison shows why. Rack-2 and rack-3 inlet sensors at one tolerance score 2.7633 and 2.7616 by spread, almost the same, and 0.8277 and 0.2821 K by need. Spread counts the fraction it narrows; only need asks whether anything depends on it.

### 11.6.2 The recommendation

The next measurement is a permanent rack-3 inlet sensor. It takes the band from 0.8283 to 0.0861 K when averaged, and 0.2821 K on one window; a repeat survey of $\beta_3$ at 0.2605 K is the one-off that competes with a single window, and it leaves no reading behind. With a rack-3 sensor, on one window, with the means held at the present fit and zero-mean error at the node, the crossings move to 148.73 kW for the loop and 220.91 kW for the frozen plant, and $\beta_3$'s spread falls to 0.0314. A real reading would also move $\beta_3$, whose truth lies +0.88 posterior standard deviations from the survey.

### 11.6.3 What the analysis showed

The sheet said a margin covered the loss and manual was the safe way to carry on. The analysis found that the sheet's question has two different plants in it, which with the airflow pinned respond to the same rise in load with opposite signs. Each therefore has its own side on which it breaks: manual at the last supply fails above 205 kW, while a hold reset to 11.87°C holds the range. The loop had added nothing to the band at the window; it had only moved the need, taking the coils out of the inlet's gradient and leaving the band on a survey's fraction no reading holds. A virtual return from the discharges and the meter is narrower than the sensor it replaces, and keeping it makes the loop depend on those harmless sensors. The one measurement that changes the picture is at rack 3, where nothing stands today.

### 11.6.4 Limits of the model

**Steady state and stability.** Every node is a steady operating point. **The stability of a loop closed on a virtual reading is not computed.** The discharges, the meter and the

integrator have dynamics that Chapter 8's method would read, and this chapter makes no claim about them. The hall's capacity is left out.

**One window and one kind of loss.** The chapter reads one window. The loss is a dropout or a stuck reading; other failures are not run, and no failure probabilities enter. A node's fresh error is a choice, tested in Section 11.3.4; the crossing it gives assumes zero mean at the node.

**Stated choices.** The set point $r = 30.33°\mathrm{C}$ and the $0.3\,\mathrm{K}$ prior on $b_r$ are stated choices. Percentages are this example's. The crossings carry $\hat{\beta}_3$ at the survey, and Section 11.4.4 shows how far its tolerance moves them.

**The signs.** The two slopes of equation (11.17) assume pinned airflow. Return-air control is not wrong in general: a plant whose airflow tracks its load holds a constant drop, and its closed slope is 0.

**The transfer and the aggregate.** The bump at the switch is left to Chapter 27. The aggregate is one map, of many a parent might be sent.

**The constructed record.** The readings are constructed from a stated truth, so the agreement tests the method on its own model, not the model against a real hall.

## 11.7 What to retain

**Definition 11.1** (Observability for a question)**.** An entry of the state is observable for a question $f$ when the readings reduce the spread on it and its gradient in $f$ is not zero. The first is read from the rank of the reading rows and the ratio $\sqrt{\Sigma_{ii}}$ over the prior; the second from $\mathbf{g}_i\sqrt{\Sigma_{ii}}$; the band itself is $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$, with its cross terms.

Three conclusions carry beyond this hall, as statements of form. Closing a loop on a reading turns the reading's error into a plant offset, and at the prior of the reading's tolerance that offset is the reading's row, so a loop adds no information at the window it closes on. A loop changes what a controlled plant asks of its sensors, since a row it holds can remove whole entries from a protected quantity's gradient. And after a loop loses its reading, the two plants that remain, at pinned airflow, differ in the signs of their response, not just in the size of their bands, so each needs its own check against the limit, with its stops in view.

Table 11.17 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 11.17: The terms of the chapter and where each is defined. Units for the two-unit hall.

| Term | Expression | Units | Meaning | Defined in |
|---|---|---|---|---|
| Entry term | $\mathbf{g}_i\sqrt{\Sigma_{ii}}$, eq. (11.1) | K | what an entry adds to a band | 11.1.1 |
| Controller's row | $T_r + b_r = r$, eq. (11.3) | K | holds the return reading | 11.1.2 |
| Reduced state | eq. (11.4) | kW, –, K | seven free entries | 11.1.2 |
| Rack inlet, closed | eq. (11.6) | °C | no coil in it | 11.1.2 |
| Virtual reading $v$, air balance | eq. (11.7) | °C | the return from other sensors | 11.1.2 |
| Frozen node, closed node | eq. (11.8), eq. (11.9) | °C | the two plants left | 11.1.2 |
| Row identity | eq. (11.15) | – | offset's prior is the reading's row | 11.2.3 |
| Closed and open gradients | eq. (11.16) | K per entry | the coils have left | 11.3.2 |
| Slopes in the load | eq. (11.17) | K/kW | opposite signs, airflow pinned | 11.3.5 |
| Down-date, refit | $\mathbf{\Sigma}_{-k}$; the row removed and solved again | – | two routes, never mixed | 11.3.6 |
| Aggregate map | $\mathbf{a}(\mathbf{z})$, eq. (11.18) | kW, kW/K | what a parent receives | 11.3.7 |
| Need and spread scores | Section 11.6.1 | K, – | two rankings of a sensor | 11.6.1 |

This closes Volume I, which has kept the controller's law as a row handed to the plant and asked what the belief can say of it. Part IV, which opens the next volume with Chapter 12, asks where that row comes from: it derives a control law from the problem of Chapter 7, under the restriction of no limits and no learning.

## 11.8 Exercises

**Exercise 11.1.** *Formulation.* Write a cold-aisle loop's row, with an offset $b_c$, and say which entries leave $T_3$'s gradient. (The row is $T_c + b_c = r$. By the air balance, $T_r = T_c + (Q + Q_o)/\sum C$, so $T_3 = T_c + \beta_3(T_r - T_c) = r - b_c + \beta_3(Q + Q_o)/\sum C$. Both coils leave the gradient, $\mathbf{g} = (\beta_3/\sum C,\ 0,\ 0,\ 0,\ 0,\ (Q + Q_o)/\sum C,\ -1)^{\mathsf{T}}$. The sign on the load is now positive: this loop warms the rack as load rises.)

**Exercise 11.2.** *Derivation.* Show that a prior on $b_r$ at the reading's standard deviation is the reading's row. (By the controller's row $b_r = r - T_r$, and with $r = y_{T_r}$, $b_r/\sigma = (y_{T_r} - T_r)/\sigma$, which is the open graph's row with its sign changed. The objective holds the square. A supply row at a vanishing standard deviation is then the pin, so the closed fit at $y_{T_s} = 14$, sd $10^{-4}$, equals the open fit, with a largest entry difference of $3.53 \times 10^{-10}$.)

**Exercise 11.3.** *Derivation.* Derive $T_3$'s closed gradient and its value at the window. (From $T_r = r - b_r$ and the air balance, $T_c = r - b_r - (Q + Q_o)/\sum C$, and $T_3 = T_c + \beta_3(T_r - T_c)$ gives equation (11.6). Differentiating gives $(-(1 - \beta_3)/\sum C,\ 0,\ 0,\ 0,\ 0,\ (Q + Q_o)/\sum C,\ -1)^{\mathsf{T}}$, and at the MAP state $(-0.0250,\ 0,\ 0,\ 0,\ 0,\ 10.1314,\ -1)^{\mathsf{T}}$.)

**Exercise 11.4.** *Application.* Find the load at which a frozen supply just holds 0.95. (Solve $T_3^*(\hat{\mathbf{z}}; Q^*) + 1.6449\sqrt{\mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}} = 27$°C with equation (11.8) and its gradient at each trial, since the spread grows with the load. The root has mean 25.5515°C and sd 0.8806 K, at $Q^* = 205.10$ kW.)

**Exercise 11.5.** *Application.* Compare a rack-2 and a rack-3 inlet sensor of 0.3 K, by entry spread and by need, and say what the comparison shows. (By entry spread they are 2.7633

and 2.7616; by need, sd $T_3$ is 0.8277 and 0.2821 K. Each narrows one fraction nothing else reads, so spread hardly tells them apart, and only rack 3's fraction is in the protected quantity's gradient.)

# Appendix A

# Atlas of control questions

This appendix can be read independently of the chapters. Each row is a question a control engineer asks. The first column gives the question in the engineer's words, as the chapter's questions table gives it; the second the general form of its category, written out or cited by equation number; the third its mathematical form for the plant of that chapter, once the plant graph and its state $\mathbf{z}$ are fixed, in that chapter's symbols; the fourth its category and the fifth its computation, the five categories and the five computations of Section 2.3.6; and the last the section in which the question is answered, or the chapter when a whole chapter answers it. The rows of Chapter 2 are the forms of its classification, Table 2.11. Chapter 3 develops the probability rules the rows rest on and holds no questions table, so it has no rows.

| Control engineer's question | General form | Mathematical form | Category | Computation | Answered in |
|---|---|---|---|---|---|
| *Chapter 1. Controlled Plants as Graphs* | | | | | |
| What temperature is the hall running at now? | $\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$, $\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$ | $p(T \mid \mathbf{y})$, $T = z_1$ | estimation | posterior | 1.3.1 |
| What is the process gain to the load? | $\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$, $\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$ | $p(K_Q \mid \mathbf{y})$, $K_Q = \mathrm{e}^{-z_2}/G_{\mathrm{ref}}$ | estimation of a derived quantity | posterior | 1.3.1 |
| Is the hall within its 24°C limit now? | $\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$ | $\Pr(T < 24^\circ\mathrm{C} \mid \mathbf{y})$ | prediction under a threshold | posterior | 1.3.2 |
| Can the hall carry 60 kW within 27.5°C at this set point? | $\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$ | $\Pr(T_2 < 27.5^\circ\mathrm{C} \mid \mathbf{y})$, $T_2 = T_c + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-z_2}$ | prediction under a threshold | posterior | 1.3.2 |
| How far must the set point move to hold it? | $\Pr(f(\mathbf{z}; r) < f^* \mid \mathbf{y})$ at a scenario node | $\Pr(T_2(r) < 27.5^\circ\mathrm{C} \mid \mathbf{y})$, $T_2(r) = r + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-z_2}$ | prediction under a threshold | posterior | 1.3.3 |
| Which added sensor would narrow the gain? | preposterior covariance $(\mathbf{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$ | $\mathbf{\Sigma}_a$ for each candidate row | design | preposterior | 1.6.1 |
| *Chapter 2. The Language of Control Questions* | | | | | |
| What is unknown, and how well is it known? | (2.9), (2.10) | $p(z_i \mid \mathbf{y})$, $\mathbb{E}[z_i \mid \mathbf{y}]$ | estimation | posterior | 2.4.1 |
| Does a requirement hold, with the stated assurance? | (2.12) | $\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) \geq \kappa$, $f^*$ the requirement | prediction under a threshold | posterior | 2.4.2 |
| What happens in the coming run? | (2.13) | $\int \Pr(A_{\mathrm{future}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$ | prediction under a threshold | posterior | 2.4.4 |

| Control engineer's question | General form | Mathematical form | Category | Computation | Answered in |
|---|---|---|---|---|---|
| Is it rising, and when does it cross its limit? | (2.21) | $\Pr(\gamma > 0 \mid \mathbf{y})$; $\Pr(t^* \le s \mid \mathbf{y}) = \Pr(\gamma > 0,\ f_s(\mathbf{z}) \ge f^* \mid \mathbf{y})$ | prediction under a threshold for the sign and a crossing by a date; estimation for a forecast and the distribution of a crossing time | trend | 2.3.4, Chapter 19 |
| Is it the path or the sensor? | (2.14), (2.15) | $\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\ \Pr(M_j)$ | hypothesis comparison; diagnosis over a family of faults | evidence | 2.5.1, 2.5.2, 2.5.3 |
| What is the expected consequence? | (2.23) | $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}] = \int L(a, \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$ | estimation of a derived quantity | expected consequence | 2.5.4 |
| Retune or leave the tuning, and is a step test worth its slot? | (2.23), (2.16) | $\arg\min_a \mathcal{R}(a)$; EVSI | design | expected consequence; preposterior | 2.5.5, 2.6.1, 2.6.2 |
| *Chapter 4. The Anatomy of Probabilistic Inference* | | | | | |
| What is the thermal resistance of the air path? | moments, (2.10) | $p(\rho \mid \mathbf{y})$: $\hat{z}_1$, $\sqrt{\Sigma_{11}}$ | estimation | posterior | 4.2.10, 4.3.6 |
| What was the hall temperature during the test? | moments, (2.10) | $p(T_1 \mid \mathbf{y})$: $\hat{z}_2$, $\sqrt{\Sigma_{22}}$ | estimation | posterior | 4.3.6 |
| What will the hall temperature be at full load? | moments, quantiles, (2.10), (2.11) | $p(T_2 \mid \mathbf{y})$: $\hat{z}_3$, $\sqrt{\Sigma_{33}}$, credible interval | estimation | posterior | 4.3.1, 4.3.4, 4.3.6 |
| Can the hall be released for 60 kW within 28°C? | threshold, (2.12) | $\Pr(T_2 < T_{\max} \mid \mathbf{y})$, with $f = -T_2$ | prediction under a threshold | posterior | 4.3.5, 4.3.7 |
| How far must the set point move to hold it at 0.95? | threshold at a scenario node, (2.12) | $r_{95}$ from $\Pr(T_2^*(r) < T_{\max} \mid \mathbf{y}) = 0.95$ | prediction under a threshold | posterior | 4.3.8 |
| *Chapter 5. Actuators and Controllers as Graph Rows* | | | | | |
| What does closing the loop change in the graph? | moments, (2.10) | the free and pinned entries of each graph; the columns of $\mathbf{J}$ | estimation | posterior | 5.3.1 |
| What is the plant now? | moments of derived quantities, (2.10) | $p(G, K_Q, W \mid \mathbf{y})$ | estimation of a derived quantity | posterior | 5.3.2 |
| Lower the set point 1 K: what moves, closed loop and open? | moments at a scenario node, (2.10) | $p(T^*, T_c^*, u^* \mid \mathbf{y})$ at $r^* = 17$°C, against $u$ pinned | prediction | posterior | 5.3.3 |
| A load step to 45 kW: what moves, closed loop and open? | moments at a scenario node, (2.10) | the same at $Q^* = 45$ kW | prediction | posterior | 5.3.4 |

| Control engineer's question | General form | Mathematical form | Category | Computation | Answered in |
|---|---|---|---|---|---|
| Will the valve reach its stop at 17°C and 45 kW? | threshold at a scenario node, (2.12) | $\Pr(u^* \geq 1 \mid \mathbf{y})$ | prediction under a threshold | posterior | 5.3.5 |
| *Chapter 6. Protected Quantities and Chance Constraints* | | | | | |
| What does an offset the loop cannot see do to the plant? | moments, (2.10) | $p(b, g, g_w \mid \mathbf{y})$, $p(T^*, u^* \mid \mathbf{y})$ | estimation | posterior | 6.3.1 |
| What must stay inside its limit, with what probability? | chance constraint, (6.1) | $\Pr(T^*(r) \leq 27.5°\text{C} \mid \mathbf{y}) \geq 1 - \alpha$ | prediction under a threshold | posterior | 6.3.2 |
| What is the highest set point that holds it? | the same, inverted, (6.17) | $r_{95}$, with back-off $\Phi^{-1}(0.95)\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}$ | prediction under a threshold | posterior | 6.3.3 |
| Is the valve a protected quantity too? | (6.1), gradient depending on $r$ | $\Pr(u^*(r) \leq 1 \mid \mathbf{y}) \geq 1 - \alpha$ | prediction under a threshold | posterior | 6.3.4 |
| Which set points hold both? | joint chance constraint, (6.20) | $\Pr(T^* \leq 27.5°\text{C},\ u^* \leq 1 \mid \mathbf{y}) \geq 1 - \alpha$ | prediction under a threshold | posterior | 6.3.5 |
| At what load does no set point hold? | the same, over the load | the window's width against $Q^*$ | prediction under a threshold | posterior | 6.3.5 |
| *Chapter 7. The Control Problem on the Graph* | | | | | |
| What is the problem every controller is solving part of? | expected consequence over policies, (2.23) | the problem (7.1) | design | expected consequence | 7.1.1 |
| With no limits, does the band change the move? | the same, restricted, (7.3) | $u_k = K_k \hat{x}_k$, against $s_0$ | design | expected consequence | 7.3.1 |
| Can the future be fit like the past? | MAP against expectation, (7.2) | $u_0$ with the future noise left free | estimation | posterior | 7.3.2 |
| What does a response to the next reading save, under a limit? | chance constraint over policies, (6.1) | fixed, linear, free, held after every reading | design | posterior | 7.3.3 |
| When the plant is unknown, what does the band do to the move? | the recursion, (7.3) | certainty equivalent, cautious, exact | design | preposterior | 7.3.4 |
| Which later chapter is which restriction? | – | the map | – | all five | 7.3.5 |
| *Chapter 8. Stability Under Parameter Uncertainty* | | | | | |
| What does the band do to the critical gain? | moments of a derived quantity, (2.10) | $p(K_c(\mathbf{z}) \mid \mathbf{y})$, with $\rho(\mathbf{M}) = 1$ at $K_c$ | estimation of a derived quantity | posterior | 8.3.1 |
| How likely is this tuning to be stable? | threshold at a scenario node, (2.12), (6.1) | $\Pr(\rho(\mathbf{M}(K_p, \mathbf{z})) < 1 \mid \mathbf{y})$ | prediction under a threshold | posterior | 8.3.2 |
| What is the largest gain that holds 0.95? | the same, inverted, (6.17) | $K_{95}$: $\Pr(K_c > K_{95} \mid \mathbf{y}) = 0.95$ | prediction under a threshold | posterior | 8.3.3 |
| And if it must also decay? | the same, on the pair, (6.1) | $\Pr(\lvert\lambda_{\text{pair}}\rvert \leq 0.8609 \mid \mathbf{y})$ | prediction under a threshold | posterior | 8.3.4 |
| Where does the band come from? | budget of a variance, (4.54) | $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$ by entry | estimation of a derived quantity | posterior | 8.3.5 |

| Control engineer's question | General form | Mathematical form | Category | Computation | Answered in |
|---|---|---|---|---|---|
| What does instability look like with stops? | simulation at the truth | $\mathbf{x}_{k+1}$ of the nonlinear loop | prediction | – (simulation at the truth) | 8.3.6 |
| Which test would narrow it most? | next measurement, (2.16) | sd $K_c$ after each candidate | design of measurement | preposterior | 8.6 |
| *Chapter 9. The Kalman Filter as a Time-Stacked Fit* | | | | | |
| Is the filter a separate device? | moments, stacked, (9.19) | $p(T_k \mid \mathbf{y}_{60:k})$ against the recursion | estimation | posterior | 9.3.1 |
| What is the hall at the last step? | moments, (2.10) | $\hat{T}_{120\mid120}$, $P_{120\mid120}$ | estimation | posterior | 9.3.2 |
| What do later readings say about earlier steps? | moments, (9.19) | $\hat{T}_{k\mid120}$ against $\hat{T}_{k\mid k}$ | estimation | posterior | 9.3.3 |
| How fast does the band fall, and to what gain? | moments of a derived quantity, (2.10) | $\upsilon_k$, $P_{k\mid k}$, the steady limit (9.21) | estimation of a derived quantity | posterior | 9.3.4 |
| What did the hall do through the load step? | moments, (9.19) | $T_k$ as a virtual reading, $k = 81, \dots, 120$ | estimation | posterior | 9.3.5 |
| With $g$ free, does the recursion still give the marginal? | moments of a derived quantity, (9.19) | $p(g, T_{120} \mid \mathbf{y}_{60:120})$, three routes | estimation of a derived quantity | posterior | 9.3.6 |
| *Chapter 10. Feedback as Conditioning* | | | | | |
| What does one reading do to the belief? | moments, (2.10), (10.5) | $(\hat{\mathbf{z}}, \boldsymbol{\Sigma}) \to (\hat{\mathbf{z}} + \mathbf{K}\iota,\ \boldsymbol{\Sigma} - S\mathbf{K}\mathbf{K}^{\mathsf{T}})$ | estimation | posterior | 10.3.1 |
| What is known before the reading comes? | moments of a preposterior, (10.8) | $\boldsymbol{\Sigma}_n$ against $n$ | estimation | preposterior | 10.3.2 |
| What can more hall readings not remove? | the same, in the limit, (10.9) | $\lim_n \mathrm{sd}\, T^*$ | estimation | preposterior | 10.3.3 |
| What does one reading do to the window? | threshold at a scenario node, (2.12), (6.17) | $r_{95}$ and the window after $y$ | prediction under a threshold | posterior | 10.3.4 |
| Should the PI row act on the reading or on a filtered estimate? | expected consequence, one fixed law, (2.23) | travel, $T_c$, margin against $L$ | design | expected consequence | 10.3.5 |
| What do a step response's rows add? | moments, (2.10) | $p(g_w, b \mid$ stacked rows$)$ | estimation | posterior | 10.3.6 |
| *Chapter 11. Observability and Sensor Loss* | | | | | |
| What can these readings resolve? | moments, (2.10) | rank of the reading rows; sd against prior, per entry | estimation | posterior | 11.3.1 |
| What does the controller need of them? | moments of a derived quantity, (2.10) | $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ for $T_3$, the hold node and $T_s^*$ | estimation of a derived quantity | posterior | 11.3.2 |
| The loop's reading drops: what is the loop? | moments at a scenario node, (2.10) | $T_3^*$, $T_r^*$, $T_s$ frozen, open loop | prediction | posterior | 11.3.3 |
| Can a virtual reading replace it? | moments at a scenario node, (2.10) | closed on the air balance, $T_3^*$, closed loop | prediction | posterior | 11.3.4 |
| Hold or virtual loop, as the load moves? | chance constraint, (6.1) | $\Pr(T_3^* \le 27°\mathrm{C} \mid \mathbf{y}) \ge 0.95$; $r_{95}$, $T_{s,95}$, crossing loads | prediction under a threshold | posterior | 11.3.5 |

| Control engineer's question | General form | Mathematical form | Category | Computation | Answered in |
|---|---|---|---|---|---|
| Which other reading matters most? | moments, row removed, (2.10) | the ranking by added band on $T_3$ | estimation | posterior | 11.3.6 |
| What does a parent level keep? | marginal of a derived quantity, (2.10) | $p(Q_{\text{tot}}, G_{\text{tot}} \mid \mathbf{y})$ by Schur complement, (11.18) | estimation of a derived quantity | posterior | 11.3.7 |
| Which sensor next? | preposterior variance, (2.16), (4.73) | ranking by $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ against by entry spread | sensor placement | preposterior | 11.6.1 |