THE COST OF EXACT BATCH DECISIONS WITH PARTIAL EVIDENCE

Anonymous authors
Paper under double-blind review

ABSTRACT

We study how much paid evidence is needed to reproduce a fixed batch decision. An evaluator may stop before resolving every item, but its output must agree with every result vector compatible with the responses it purchased. We characterize the resulting saving through local query policies. Each policy trades a reduction in fee against uncertainty left in the batch count. Pricing this tradeoff gives an additive comparison with the optimum over all legal queries. The error separates omitted local policies, insufficient copies and threshold geometry. For a single threshold, bounded representation and supply losses give an $O(1)$ remainder without a variance lower bound. When shared response probabilities are unknown, purchased paths provide learning feedback through their fees and support endpoints. Under explicit interface and completion-cost conditions, learning adds a cost governed by the remaining query opportunities. At a critical threshold this cost is sharply of order $N^{1/4}$, against a saving of order $\sqrt{N}$. Finite calculations show both the attainable saving and the acquisition costs that can consume it.

1 INTRODUCTION

Many evaluation workloads end in a decision about a batch. A release gate compares evaluation metrics with configured thresholds (TensorFlow Authors, 2024). Code evaluation suites produce test outcomes for fixed programs (Liu et al., 2023). Given these outcomes, a batch gate can ask whether the programs meet a prescribed pass rate. The requested answer can be much smaller than the evidence needed to evaluate every item. Once enough outcomes establish acceptance or rejection, the remaining results cannot change that answer.

This suggests evaluating only as much as the decision requires. The difficulty is deciding which evidence to acquire first. A short test prefix may certify a failure, while a passing prefix leaves an expensive remainder. A cheap predictor may suggest which item to inspect, but buying that prediction also consumes resources. Its accuracy alone does not determine its value: the relevant question is whether it saves more subsequent work than it costs. The same partial observation may be valuable far from a decision boundary and almost useless when one unread result can change the answer.

We fix the items, query interface and output rule, and ask: *how much evidence can be omitted while reproducing exactly the decision of full evaluation?* Exactness concerns this realized batch. Every completion compatible with the purchased responses must give the same answer. A response model helps choose economical queries. The final certificate uses what the responses logically establish. This separation permits statistical learning of query choices while preserving the evaluator's prescribed decision on every batch.

The analysis must connect three levels. A local query changes one item's possible outcomes and expected fee. A batch threshold determines how much unresolved evidence the final decision can tolerate. An adaptive policy must turn these local opportunities into a certificate using the items actually available. We express the connection through two uncertainty budgets, one for each endpoint of the attainable count interval. A local policy spends these budgets and saves a fee. Their prices therefore predict the value of leaving evidence unread.

Our first result compares that prediction with the optimum over all legal batch queries. It separates the losses from missing local policies, insufficient copies of useful item types, and category boundaries.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

When these losses are bounded, the predicted saving has only an additive $O(1)$ error. This identifies when a compact local description suffices to price a large heterogeneous batch.

Our second result asks how much of that saving survives when shared response probabilities must be learned. The learner observes only the paths it buys. Their fees and attainable endpoints provide feedback even when final item results remain unread. Near a threshold on the scale of ordinary count fluctuations, the saving is of order $\sqrt{N}$ and the additional learning fee is sharply of order $N^{1/4}$ under the stated conditions. Finite calculations then test the economic meaning of this separation. They show both a benefit from informative queries and regimes where learning fees consume it. The paper proceeds from the exact output contract to local prices, batch costs, and learning those same prices.

2 Related work

Exact evaluation and certification. Sequential testing chooses observations until they determine a fixed function (Ünlüyurt, 2004). Optimal testing rules are known for k-out-of-n systems (Salloum and Breuer, 1984; Chang et al., 1990). Stochastic Boolean evaluation and score classification extend the target to more general functions and score intervals (Deshpande et al., 2014; Gkenosis et al., 2018; 2022; Plank and Schewior, 2024; Ghuge et al., 2025). Their cost-approximation guarantees still require a correct function value; approximation concerns the acquisition fee. Recent decision-diagram methods optimize expected cost for general formulas under a supplied joint law (Zong et al., 2026). One-sided certification with correlated observations is studied through a joint-distribution oracle (Ghuge et al., 2026). Our target belongs to this exact-evaluation tradition. We study an additive cost description for independent components with rich local query paths, and make its representation, supply and boundary errors explicit.

Paying for information. Pandora's problem prices inspection against the value of the revealed alternative (Weitzman, 1979). Markovian information acquisition allows several paid steps before an item's reward is resolved (Gupta et al., 2019). These models establish the importance of charging the path to useful information. Here a local path saves a full-evaluation fee but leaves uncertainty in a count whose category must be certified. The resulting price measures saved fee per unit of residual uncertainty. It couples local acquisition choices to the batch's stopping interval.

Efficient model evaluation. Active Testing chooses labels to reduce evaluation-estimator variance while correcting acquisition bias (Kossen et al., 2021). Active Evaluation Acquisition predicts unqueried benchmark outcomes using learned dependencies across examples (Li et al., 2025). Dynamic feature acquisition similarly trades observations for predictive utility (Covert et al., 2023; Guney et al., 2025). Learn then Test and prediction-powered inference attach statistical guarantees to calibration and estimation (Angelopoulos et al., 2025; 2023). Our object is the category of a fixed realized batch. Even a highly probable result must be queried if its alternative can change that category. This connects acquisition cost to logical certificates rather than estimator precision.

Learning which queries to buy. Product-distribution learning gives sample bounds for near-optimal policies, including Pandora's problem (Guo et al., 2021). Bandit Pandora learns from policy-value feedback over fresh rounds (Gatmiry et al., 2024). Fixed-confidence best-arm identification learns which unknown mean is largest (Garivier and Kaufmann, 2016). Here repeated local interfaces share unknown response probabilities. Learning and certification spend queries in the same batch, and the available learning opportunities end when its decision is known. The comparison therefore charges both exploration and final completion. Its feedback is the executed path's fee and support endpoints.

The analysis joins local decision processes to constrained allocation. Occupation measures express local query costs (Altman, 1996). Bounded-loss allocation methods motivate adjusting allocations as resources change (Arlotto and Gurvich, 2019; Bumpensanti and Wang, 2020; Vera and Banerjee, 2021). Episodic policy learning handles known (Zimin and Neu, 2013) and unknown transitions (Jin et al., 2020). We connect these tools to exact certification through the uncertainty budgets developed below.

3 Problem setup

The distinction between estimation and exact evaluation is a stopping rule. We first make that rule concrete, then specify the evidence and fees available to an evaluator. Consider five fixed programs.

A gate accepts when at least three pass. A prefix test costs one. A failed prefix certifies failure; a passing prefix permits a remainder test of fee nine. After three complete passes, the number of passing programs lies in $[3, 5]$. Both unread results may change without changing the gate. On a batch of five passing programs, this path costs 30 instead of the full-evaluation fee 50. A proxy may change the query order, but its prediction alone cannot certify a pass.

The output and its certificate. A batch has N components. Component g has a hidden integer contribution $X_g \in \{0, \dots, x_{\max,g}\}$. The binary program example has $X_g = 1$ exactly when all its tests pass. Set

$$K = \sum_{g=1}^{N} X_g, \qquad n = \sum_{g=1}^{N} x_{\max,g}.$$

A threshold k asks whether $K \geq k$. Distinct thresholds $1 \leq k_1 < \cdots < k_q \leq n$ define a category map Γ with cells $[0, k_1 - 1], [k_1, k_2 - 1], \dots, [k_q, n]$. Some interfaces also require a local action A_g for each component. When only the category is needed, each A_g is constant.

Let $\mathfrak{h}_g$ contain the responses purchased on component g. Among all local patterns compatible with that history, let l_g and u_g be the smallest and largest possible contribution. Then

$$L = \sum_{g} l_g, \qquad U = \sum_{g} u_g, \qquad L \leq K \leq U.$$

For a single threshold the stopping test is

$$L \geq k \implies \Gamma(K) = \text{accept}, \qquad U < k \implies \Gamma(K) = \text{reject}.$$

For several thresholds, $[L, U]$ must lie in one category cell. All required actions must also be known. This is *support-exact* evaluation: the output $(A_1, \dots, A_N, \Gamma(K))$ is identical on every compatible completion. It does not depend on a confidence level.

The paid query interface. A local mechanism $\omega_g \in \Omega_g$ specifies every query response, $A_g(\omega_g)$ and $X_g(\omega_g)$. The finite support, legal query graph, prerequisites and positive fees are known initially. Responses are hidden until purchased, and each address is charged once. Every history admits a legal full evaluation of (A_g, X_g). For the program example, a proxy H, prefix bit P and complete-suite bit X satisfy $X \leq P$. Their fees are c_h, c_p, r, with the remainder available after $P = 1$. The prefix–remainder routine has expected fee $c_p + r \Pr(P = 1)$. Its order is optimal for every positive law on this support. The batch law is $\mu = \bigotimes_g \mu_g$, where each μ_g is positive on its declared support. Components have disjoint query addresses. Responses within a component may be dependent. A repeated type has the same graph, targets, fees and law within the batch. This permits shared probability estimates across its copies, while the result of one copy never certifies another's hidden outcome. Probabilities are known for the cost comparison and unknown for learning. The legal support is known in both.

The cost to be minimized. A policy Π selects legal queries from purchased responses and private randomization. Write $\mathfrak{h}_g^0$ for the initial history and let $v_g(\mathfrak{h})$ be the least expected fee for fully determining (A_g, X_g) from $\mathfrak{h}$. Full evaluation and optimal batch evaluation have expected fees

$$V_0 = \sum_{g} v_g(\mathfrak{h}_g^0), \qquad C_N^* = \inf_{\Pi \text{ support-exact}} \mathbb{E} C_\Pi.$$

Here C_Π is the total fee of purchased queries. The saving is $V_0 - \mathbb{E} C_\Pi$, so all proxy, access and completion fees are charged before a saving is reported. The oracle defining C_N^* knows the laws and may use every legal query. It does not see an unpurchased response. The comparison concerns query fees; planning time is a separate cost.

4 Pricing the evidence left unread

The interval $[L, U]$ certifies when to stop, but does not yet say which queries are worth buying. To compare them, we measure the expected fee saved and the unresolved uncertainty they leave. Let $\Omega_g(\mathfrak{h})$ be the patterns compatible with a local history, and define

$$m_g(\mathfrak{h}) = \mathbb{E}[X_g \mid \mathfrak{h}], \qquad b_g^+(\mathfrak{h}) = m_g(\mathfrak{h}) - l_g(\mathfrak{h}), \qquad b_g^-(\mathfrak{h}) = u_g(\mathfrak{h}) - m_g(\mathfrak{h}).$$

Table 1: **Hints trade fee savings for lower uncertainty.** The full fee is r. A hint lowers the saving per item, but reduces one directional exposure from $1/2$ to $\varepsilon_h/2$. The selective policies exchange B^+ and B^-; their value is determined by the gain-to-exposure ratio. Full evaluation has zero gain and exposure.

Local policy	$\mathbb{E}C_\pi$	B^+_π	B^-_π	G_π
Leave Y unread	0	$1/2$	$1/2$	r
Buy H; read Y only if $H=1$	$c_h + r/2$	$\varepsilon_h/2$	$(1-\varepsilon_h)/2$	$r/2 - c_h$
Buy H; read Y only if $H=0$	$c_h + r/2$	$(1-\varepsilon_h)/2$	$\varepsilon_h/2$	$r/2 - c_h$

The two quantities are the distances from the conditional mean to the attainable endpoints. They quantify residual uncertainty in the count. They do not change the support-exact stopping test.

A local partial policy π stops once its action is known, possibly leaving its contribution unresolved. Its paid fee C_π and terminal history $\mathfrak{h}_\pi$ define a *net column*:

$$B^\pm_{g\pi} = \mathbb{E}b^\pm_g(\mathfrak{h}_\pi), \qquad G_{g\pi} = v_g(\mathfrak{h}^0_g) - \mathbb{E}C_\pi. \tag{1}$$

Thus $(B^+_{g\pi}, B^-_{g\pi}, G_{g\pi})$ records two expected uncertainty costs and one expected fee saving. Acquisition on every branch enters $G_{g\pi}$. A positive gain has positive exposure in both directions. We suppress g only when the local type, law and fees are fixed.

For a running calculation, let Y be a fair bit with a known action. Reading Y costs r. A hint $H = Y \oplus E$ costs c_h, where E is independent Bernoulli(ε_h), $0 < \varepsilon_h \le 1/2$. Table 1 compares leaving Y unread with using the hint selectively. At $(r, c_h, \varepsilon_h) = (10, 1, 0.1)$, retaining only the $H = 0$ branch gives

$$\mathbb{E}C_\pi = 1 + 10\Pr(H=1) = 6, \qquad (B^+_\pi, B^-_\pi, G_\pi) = (0.05, 0.45, 4).$$

It saves less per item than leaving the bit untouched, but uses much less lower-side uncertainty. Its ratio $G_\pi/B^+_\pi = 80$ exceeds the untouched ratio $10/(1/2) = 20$. This is the reason to buy the hint.

At a joint record, write $M = \sum_g m_g$. For a category cell $[a, b]$, the stopping condition becomes

$$[L, U] \subseteq [a, b] \quad \Longleftrightarrow \quad \sum_g b^+_g = M - L \le M - a, \qquad \sum_g b^-_g = U - M \le b - M.$$

These are the two uncertainty budgets. A library $\mathcal{C}$ of positive-gain local policies turns them into a fee-saving allocation:

$$\begin{aligned} \mathcal{V}_0(c) &= \max_{z \ge 0} \left\{ \sum_{\pi \in \mathcal{C}} G_\pi z_\pi : \sum_{\pi \in \mathcal{C}} B^s_\pi z_\pi \le c_s \ (s = +, -) \right\}, \\ \mathcal{D}_0 &= \{\lambda \ge 0 : \lambda \cdot B_\pi \ge G_\pi \ (\pi \in \mathcal{C})\}, \qquad \mathcal{V}_0(c) = \min_{\lambda \in \mathcal{D}_0} \lambda \cdot c. \end{aligned} \tag{2}$$

Here $c = (c_+, c_-)$ is the budget and z_π the fractional number of copies assigned to policy π. The nonnegative vector λ prices the two exposures. This relaxation permits unlimited copies; a real controller must use the batch's actual supply.

For $x \in [a, b]$, put $F_0(x) = \mathcal{V}_0(x - a, b - x)$. Omit the lower budget at $a = 0$ and the upper at $b = n$, and set $F_0 = 0$ on unit gaps between cells. The predicted saving and axial prices are

$$S_{0,N} = \mathbb{E}F_0(K), \qquad \widehat{\rho}_{0,s} = \max_{\pi \in \mathcal{C}} G_\pi / B^s_\pi, \qquad \Lambda_0 = \widehat{\rho}_{0,+} + \widehat{\rho}_{0,-}. \tag{3}$$

With one threshold, only one budget is present on each side:

$$S_{0,N} = \widehat{\rho}_{0,+}\mathbb{E}(K - k)_+ + \widehat{\rho}_{0,-}\mathbb{E}(k - 1 - K)_+. \tag{4}$$

Here $(x)_+ = \max\{x, 0\}$. The distance from the count to the decision boundary determines how much evidence can remain unread; the local price determines the fee saved by that uncertainty. Using all tagged positive policies gives the full-batch objects $\mathcal{V}, \mathcal{D}, \mathcal{F}^{\mathrm{net}}$ and S_N.

In the hint example, the columns in Table 1 give

$$\widehat{\rho}_+ = \widehat{\rho}_- = \max\{2r, (r - 2c_h)/\varepsilon_h\}. \tag{5}$$

The hint improves the price exactly when $c_h < r(1/2 - \varepsilon_h)$. Accuracy alone therefore does not determine its value. At the running parameters the price increases fourfold, from 20 to 80; actual finite savings also depend on supply and completion. The next result quantifies those losses.

5 From local prices to batch cost

Local prices predict a saving as if useful policies could be copied freely. The next theorem asks how accurately they describe a real batch. Its lower bound tests every legal query policy. Its upper bound gives a policy that uses the available items and pays to finish its certificate. Three corrections make this comparison precise.

First, the library may omit a useful way to query an item. Select at most M_{lib} policies from J repeated types. These types form the *core*; the remaining items form the exterior. At price λ, let $U_{g,\lambda}(\mathfrak{h}_g^0)$ minimize $\mathbb{E}C_\pi + \lambda \cdot B_{g\pi}$ over all legal local partial policies, including full evaluation. This is a backward query calculation with an extra charge for stopping with unresolved evidence. Let $\mathcal{P}_0$ be the vertices, or corners, of the library price region $\mathcal{D}_0$. The largest local shortfall at these prices is

$$d_g = \max_{\lambda \in \mathcal{P}_0} [v_g(\mathfrak{h}_g^0) - U_{g,\lambda}(\mathfrak{h}_g^0)]. \tag{6}$$

Here $0 \le d_g \le v_g(\mathfrak{h}_g^0)$; it is zero when the library prices already cover every local policy. The lower comparison prices an item using the other items' outcomes. For this step, an item can belong to a set $\mathcal{A}$ if a fixed reference value remains inside every unresolved interval:

$$\exists a_g^{\mathrm{anc}} \in \bigcap_{\mathfrak{h}:\ A_g \text{ known},\ X_g \text{ unresolved}} [l_g(\mathfrak{h}), u_g(\mathfrak{h})]. \tag{SA}$$

An unresolved binary result always has interval $[0, 1]$, so it satisfies this condition. Charge the full fee for components outside $\mathcal{A}$. The total representation loss is

$$\mathfrak{D}_N = \sum_{g \in \mathcal{A}} d_g + \sum_{g \notin \mathcal{A}} v_g(\mathfrak{h}_g^0). \tag{7}$$

Thus a few exceptional interfaces can be handled through their total fee.

Second, several categories can make the useful uncertainty budget change as the count crosses an internal cell. Write $x_{\max} = \max_g x_{\max,g}$. Let w_Γ be the largest internal-cell width, zero with one threshold, and let $H_\Gamma = k_q - k_1$ be the total threshold span. Define

$$\mathfrak{p}_N = \sup_z \Pr(K = z), \quad \zeta_N = \min\{w_\Gamma, \mathfrak{p}_N H_\Gamma\}, \quad \tau_N = 1 + \sigma_N + d_0, \tag{8}$$

where $\sigma_N^2 = \operatorname{Var} K$ and $d_0 = d_\Gamma(\mu_N)$, $\mu_N = \mathbb{E}K$. Here $d_\Gamma(x)$ is the distance to the nearest present cell boundary, and is zero in a unit gap. The correction ζ_N measures category geometry; it is zero for one threshold. The scale τ_N combines count fluctuations and distance from the mean to a boundary. It sets the size of the useful reserve.

Third, the allocation must fit the physical supply and admit affordable completion. With two budgets, an allocation basis uses at most two policies with positive coefficients. Impose the following local bounds. Let $\beta_{\mathrm{gen}} = \min_{\pi \in \mathcal{C}, s} B_\pi^s > 0$. Every selected nonsingular two-column exposure matrix Q satisfies $\|Q^{-1}\|_\infty \le \kappa$. This controls how much an allocation changes when its budgets change. Core partial and full paths cost at most C_{loc}. For each direction s, legal terminal completions have conditional expected fee at most $\chi_{\mathrm{rec}} b_g^s(\mathfrak{h})$. A one-column basis uses $1/B_\pi^s \le 1/\beta_{\mathrm{gen}}$. For a sequence of batches these bounds, J, M_{lib} and $x_{\max}$ are uniform, while the laws, fees and unused portions of the graphs may vary.

To describe actual supply, reserve m_0 untouched copies per library type and fully evaluate the exterior. Let M_0 be the resulting posterior count mean, $a_R = 8/\beta_{\mathrm{gen}}$ and $b_R = 1 + a_R J x_{\max}$. If the reserve is too small, the policy completes it fully. The predicted saving lost on this branch is

$$\mathcal{L}_N(m_0) = \mathbb{E}[F_0(K)\mathbf{1}\{m_0 \le a_R d_\Gamma(M_0) + b_R\}].$$

We can now compare the prediction $V_0 - S_{0,N}$ with both the optimum and the implemented policy. Each side of the bound has a different role: the shortfall controls omitted opportunities; the supply loss controls whether the chosen opportunities can be executed.

Theorem 1 (Cost with a known local library)**.** *Suppose the core laws and optimal full routines are available. Under the product query model and the displayed local execution bounds, choose any finite threshold family, any anchored set $\mathcal{A}$, and any available $m_0 \geq 1$ copies per library type. There are constants $c_{\mathrm{low}}, c_{\mathrm{up}}$ depending only on those local bounds and a support-exact controller such that*

$$\begin{aligned} V_0 - S_{0,N} - \mathfrak{D}_N - c_{\mathrm{low}}\zeta_N \leq C_N^* \leq \mathbb{E}C_{\mathsf{Track}_0} \\ \leq V_0 - S_{0,N} + \mathcal{L}_N(m_0) + c_{\mathrm{up}}(1+\zeta_N). \end{aligned} \tag{9}$$

For one threshold, bounded $\mathfrak{D}_N$ and $\mathcal{L}_N$ give $C_N^* = V_0 - S_{0,N} + O(1)$ wherever the threshold lies. For several categories, ζ_N records the additional geometric loss. The controller uses core probabilities, supplied full routines and purchased exterior outcomes.

The shortage loss has a variance-dependent certificate. Bernstein's inequality in Lemma D shows that

$$m_0 \geq a_R \left[d_0 + 2\sigma_N \sqrt{\log(2+\tau_N)} + \tfrac{4}{3} x_{\max} \log(2+\tau_N) \right] + b_R \tag{10}$$

makes $\mathcal{L}_N(m_0) \leq 2\Lambda_0$. The *critical regime* has count spread $\sigma_N = \Theta(\sqrt{N})$ and boundary distance $d_0 = O(\sqrt{N})$. The certificate requires order $\sqrt{N \log N}$ copies, attained in the worst case by the two-fee example in Appendix C.4. The finite condition for a positive certified saving is

$$S_{0,N} > \mathcal{L}_N(m_0) + c_{\mathrm{up}}(1+\zeta_N).$$

It compares the value of partial evidence directly with the fees needed to realize it.

The lower comparison applies to every legal batch policy. For any local policy and library vertex λ, the definition of d_g gives

$$\mathbb{E}C_\pi + \lambda \cdot B_{g\pi} \geq U_{g,\lambda}(\mathfrak{h}_g^0) \geq v_g(\mathfrak{h}_g^0) - d_g, \qquad G_{g\pi} \leq \lambda \cdot B_{g\pi} + d_g.$$

Freezing the other mechanisms makes a global policy a legal local policy on g. Summing these inequalities accounts for the library shortfall. The certified count cell supplies the uncertainty budgets. Posterior centering controls the remaining boundary error by ζ_N.

The upper comparison is an executable allocation. The controller first evaluates the exterior and holds a reserve of each library type. It fully evaluates reserve items in balanced groups until the remaining copies and current mean satisfy

$$a_R d_\Gamma(M) < m \leq a_R d_\Gamma(M) + b_R, \qquad \text{copies required per type} \leq \frac{d_\Gamma(M)}{\beta_{\mathrm{gen}}} < \frac{m}{8}.$$

It then executes partial policies and updates the allocation from the observed endpoints. Small allocations are activated only after they are large enough to track. Legal completion queries remove any remaining uncertainty until $[L, U]$ lies in one cell. This last step preserves exactness on every path; its fee is included in the upper bound. The complete projection, tracking and completion argument appears in Appendix C.4.

6 Learning the evidence prices

The first theorem prices a library using known response probabilities. We now keep the same query graph and exact stopping rule, but learn those probabilities from repeated types within the batch. Each exploratory query has two effects: it informs future choices and consumes part of the opportunity to save. The benchmark remains the known-law optimum C_N^*.

A retained item's result is unread, so its residual mean cannot serve as observed feedback. That mean cancels when comparing policies of the same type. For a price λ fixed before local execution,

$$\mathbb{E}C_\pi + \lambda \cdot B_\pi = \mathbb{E}[C_\pi - \lambda_+ l_g(\mathfrak{h}_\pi) + \lambda_- u_g(\mathfrak{h}_\pi)] + (\lambda_+ - \lambda_-)\mathbb{E}X_g. \tag{11}$$

The final term is common to policies of the same type. Thus the observed fee and support endpoints compare policies even when their retained results are never read. This is the feedback connecting learning to the cost characterization.

Learning must also budget for certifying the last unresolved items. Choose legal completions $\kappa_j(\mathfrak{h})$ from the known graph and fees, with fixed tie breaking. The given number χ_{rec} satisfies

$$\mathbb{E}_\mu[C_{\kappa_j(\mathfrak{h})} \mid \mathfrak{h}] \le \chi_{\mathrm{rec}} b_j^s(\mathfrak{h}) \quad (s = +, -) \tag{12}$$

at action-known unresolved core histories. The number χ_{rec} bounds the fee for removing one unit of residual uncertainty. A known floor $b_j^s \ge b_0 > 0$ and path fee at most C_{loc} suffice with $\chi_{\mathrm{rec}} = C_{\mathrm{loc}}/b_0$. The displayed cost condition also permits smaller exposures when their completion fees decrease proportionately.

The learner also receives full-evaluation tables $\boldsymbol{\sigma} = (\sigma_g)_{g \le N}$ satisfying $\mathbb{E}_{\mu_g} C_{\sigma_g} = v_g(\mathfrak{h}_g^0)$, but not their expected fees. These tables are fixed before the batch is drawn. Their identities are public information about the law, so a learning instance is the pair $(\mu, \boldsymbol{\sigma})$. The code interface supplies the same optimal prefix–remainder order for every law. Lower bounds similarly use a common routine tuple. Any setup fees for the supplied tables or numerical bounds are added to the batch fee.

For a given $\beta_0 > 0$, define the comparison library to contain all core policies with $G_\pi > 0$ and $\min_s B_\pi^s \ge \beta_0$. It is assumed nonempty, but need not generate the full core or batch region. Its true F_0 and $\mathfrak{D}_N$ are analysis quantities; the learner does not receive the identities of these policies.

Given $2s$ copies of each core type, the learner completes s pilot components and estimates their means by $\widehat{p}_j$. It holds the other s copies in reserve. With m reserve copies per type and completed outside count O, it uses $\widetilde{M} = O + m\sum_j \widehat{p}_j$ in the balanced reserve rule. Set $a_L = 8/\beta_0$, $b_L = 1 + a_L J x_{\max}$, and let $\widetilde{M}_0$ be this estimate before the reserve is reduced. The value lost on the initial full-completion branch is

$$\widetilde{\mathcal{L}}_N(s) = \mathbb{E}[F_0(K)\mathbf{1}\{s \le a_L d_\Gamma(\widetilde{M}_0) + b_L\}].$$

On the remaining branch, the learner compares observed path scores on fresh reserve items, updates prices from endpoint deficits, and completes the certificate. The next bound separates library and geometry errors from the extra cost of learning. More pilots improve mean estimates; a larger useful reserve creates more opportunities to learn and spend.

Theorem 2 (Learning an imperfect library)**.** *In the product model, suppose J declared core types have known finite query graphs, path fees bounded by C_{loc}, $2s$ independent copies per type, and the given completion-price and covered-library conditions above. Laws and fees may vary between batches under these uniform bounds. For every joint instance with the supplied optimal full routines, one support-exact executed-path algorithm satisfies*

$$\mathbb{E}_\mu C_{\mathsf{Learn}(\boldsymbol{\sigma})} - C_N^*(\mu) \le \mathfrak{D}_N + c_{\mathrm{low}}\zeta_N + c_{\mathrm{learn}}\left(\sqrt{\tau_N} + \frac{\tau_N}{\sqrt{s}}\right) + \widetilde{\mathcal{L}}_N(s). \tag{13}$$

The constants are uniform in the unknown laws and public routine tuple; they depend on $J, x_{\max}, C_{\mathrm{loc}}, \chi_{\mathrm{rec}}, \beta_0$ and the explicit core graph sizes. The quantity $\widetilde{\mathcal{L}}_N(s)$ is the library value lost on the initial shortage branch defined above. The learner uses all legal graph policies; it need not invert a library basis.

The two learning terms have different sources. The term $\sqrt{\tau_N}$ pays for choosing policies and completing the certificate on a reserve of expected size $O(\tau_N)$. The term $\tau_N/\sqrt{s}$ pays for estimating where that reserve should start. The selected reserve depends on the pilots. If e is the sum of the absolute core-mean errors,

$$\mathbb{E}m^2 = O(\tau_N^2), \qquad \mathbb{E}e^2 = O(s^{-1}), \qquad \mathbb{E}[me] \le \sqrt{\mathbb{E}m^2\mathbb{E}e^2} = O(\tau_N/\sqrt{s}). \tag{14}$$

Executed-path learning and completion each cost $O(\sqrt{\mathbb{E}m})$. The initial pilot-based count estimate has variance at most $2\sigma_N^2$. Bernstein's inequality then bounds the shortage loss when $s \ge \tau_N$ and s exceeds a sufficiently large multiple of $d_0 + \sigma_N\sqrt{\log(2+\tau_N)} + \log(2+\tau_N)$. The learning fee is then $O(\sqrt{\tau_N})$. Execution uses s and the observed means.

With $\tau_N = \Theta(\sqrt{N})$ and bounded representation and geometric errors, this yields $O(N^{1/4})$ excess cost. It is sharp on the common- routine two-sensor family of Corollary 1: the hint noises are $1/4 - \theta$

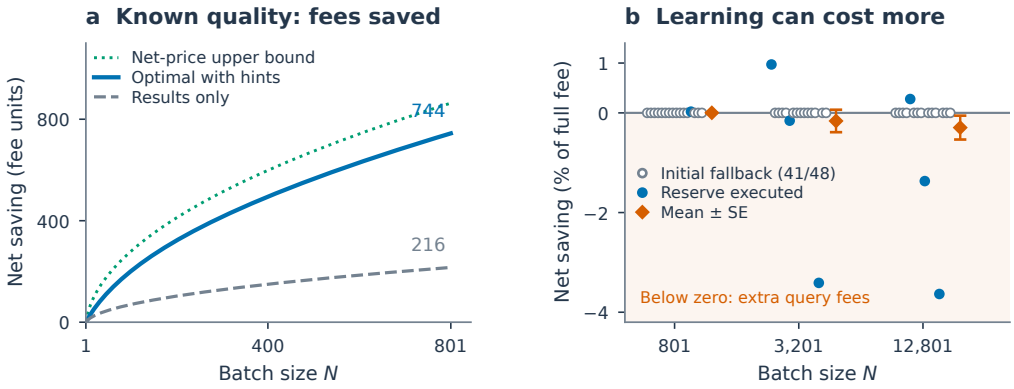

Figure 1: **Evidence value and the cost of learning it.** Both panels use the same fair-bit hint interface and majority gate; all acquisition fees are included. **a**: With known quality, hints increase the finite optimal saving beyond exact result-only stopping. The 401 archived odd-N DP values are bounded above by the net-price curve S_N (dotted). **b**: Net savings of the unknown-law procedure after pilot, hint and completion fees. All 16 batches per size are shown; diamonds are means $\pm$ one standard error. Negative values mean extra fees. Panel a reports fee units; panel b reports percentages of the full fee.

and $1/4 + \theta$, and all original query orders remain legal. A first hint bought before its result is one learning opportunity. The lower bound accounts for their data-dependent count and charges the complete query transcript. Whenever the oracle saving $\Delta_N^* = V_0 - C_N^*$ is at least $s_0\sqrt{N}$ with $s_0 > 0$,

$$\frac{V_0 - \mathbb{E}C_{\mathsf{Learn}}}{\Delta_N^*} = 1 - O(N^{-1/4}) \tag{15}$$

under these conditions.

7 What the finite costs show

The two theorems separate evidence value from the fees needed to exploit it. We now examine that separation at finite sizes. The first calculation compares useful hints with the cost of learning to use them. The second holds the evidence fixed and changes the gate.

Evidence value and the fee for learning it. First consider the fair-bit hint model from Table 1, with $(r, c_h, \varepsilon_h) = (10, 1, 0.1)$ and majority threshold $(N+1)/2$. Figure 1a asks how much the informative query adds beyond exact stopping with result queries alone. At $N = 801$, full evaluation costs 8010. The optimal known-law policy saves 744.24, versus 215.89 using only result queries. These optima come from a dynamic program over all useful legal queries. The net-price curve is their saving upper bound. The generic sufficient reserve in (10) is $8402 > N$ here, so the DP supplies the finite comparison.

Figure 1b measures the finite query fee of learning on the same interface, as in Theorem 2. Charging every pilot, hint and completion, the 16 independent batches per size give mean savings $0.125, -51.94, -378.00$ at $N = 801, 3201, 12801$. All 48 decisions are exact; 41 batches take the initial full-completion branch. On the other branches, net saving equals the avoided result fees minus the purchased hint fees. Appendix E.2 gives this pathwise accounting.

The same evidence can be cheap or expensive to certify. The second figure isolates the role of the decision boundary. A query discloses one immutable test bit from a fixed code-result record. The baseline already stops each program at its first failing test. On 148 Qwen and 144 Llama programs, we replay every threshold and both prescribed program orders, giving 584 threshold–order cases.

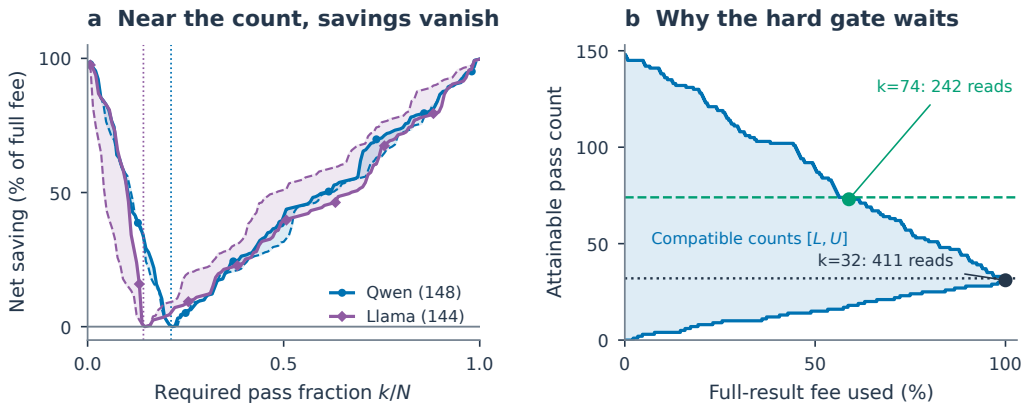

Figure 2: **The same evidence can be cheap or expensive to certify. a**: Savings across thresholds on fixed code batches. Solid/dashed curves use ascending/reversed program order; shading shows their range. Dotted lines separate acceptance and rejection. **b**: The compatible-count interval on one Qwen path. The gate $k = 74$ rejects early; $k = 32$ requires every baseline disclosure. Both orders have zero saving near the realized count. Fees count test-bit disclosures. The baseline stops each program at its first failing test.

Figure 2a holds the records fixed and changes only the required pass count. Savings collapse near the actual pass counts 31 and 20, and increase for gates that are easier to certify.

Figure 2b shows why, along one Qwen query sequence. The half-batch gate $k = 74$ rejects after 242 of the baseline's 411 disclosures, saving 41.12%. The gate $k = 32$ remains compatible with both answers until all 411 disclosures are bought. For Llama, the corresponding fees are 221 of 365 at $k = 72$, but all 365 at $k = 21$. These difficult gates need the entire baseline in both orders. The attainable count interval explains the difference: a distant gate can be decided while many tests remain unread; a gate adjacent to the realized count keeps both answers possible. These fixed-record fees illustrate the threshold-distance mechanism in (4).

Appendices E.1–E.2 give the DP recurrence and fixed protocols. Query fees are separate from planning: a Bellman pass is linear in the explicit graph, which can contain 3^d histories for d binary addresses.

8 Conclusion and Future Work

Exact batch evaluation can save evidence because the requested category tolerates unresolved item outcomes. Local net prices quantify that tolerance in fee units. The cost comparison accounts for omitted policies, finite supply and category geometry. Learning uses the same paid paths and preserves the same certificate. In the critical regime, its fourth-root overhead is smaller than the square-root oracle saving. The finite results show why acquisition and completion fees remain decisive at practical batch sizes.

Two directions follow from this accounting. The first is to design tighter finite reserve rules that turn the available evidence value into positive net savings at smaller batches. The second is to price the acquisition of reusable type information and completion bounds within a realistic evaluation interface. Both should be assessed against the same exact output and a full ledger of setup, exploration and completion fees.

AI USE STATEMENT

Generative AI tools assisted manuscript and proof writing, literature verification, and experimental code development. Verification used cited sources, explicit derivations, finite tests and recorded-output replay. The authors take responsibility for the final text, claims, code and results.

REPRODUCIBILITY STATEMENT

Sections 3–6 give the model and assumptions. Appendices A–D contain the proofs; Appendix E gives computational and experimental protocols. The source tree provides figure scripts, record-selection rules and finite verification programs.

ETHICS STATEMENT

The experiments use synthetic response models and public benchmark records. Applications should assess whether the evaluator's tests and thresholds provide appropriate coverage and error costs.

REFERENCES

E. Altman. Constrained Markov decision processes with total cost criteria: Occupation measures and primal LP. *Mathematical Methods of Operations Research*, 43:45–72, 1996. https://doi.org/10.1007/BF01303434.

A. N. Angelopoulos, S. Bates, E. J. Candès, M. I. Jordan, and L. Lei. Learn then Test: Calibrating Predictive Algorithms to Achieve Risk Control. *The Annals of Applied Statistics*, 19(2):1641–1662, 2025. https://doi.org/10.1214/24-AOAS1998.

A. N. Angelopoulos, S. Bates, C. Fannjiang, M. I. Jordan, and T. Zrnic. Prediction-Powered Inference. *Science*, 382(6671):669–674, 2023. https://doi.org/10.1126/science.adi6000.

A. Arlotto and I. Gurvich. Uniformly Bounded Regret in the Multisecretary Problem. *Stochastic Systems*, 9(3):231–260, 2019. https://doi.org/10.1287/stsy.2018.0028.

J.-Y. Audibert and S. Bubeck. Minimax Policies for Adversarial and Stochastic Bandits. *COLT*, 2009. https://www.cs.mcgill.ca/~colt2009/papers/022.pdf.

P. Bumpensanti and H. Wang. A Re-solving Heuristic with Uniformly Bounded Loss for Network Revenue Management. *Management Science*, 66(7):2993–3009, 2020. https://doi.org/10.1287/mnsc.2019.3365.

M.-F. Chang, W. Shi, and W. K. Fuchs. Optimal Diagnosis Procedures for k-out-of-n Structures. *IEEE Transactions on Computers*, 39(4):559–564, 1990. https://doi.org/10.1109/12.54850.

Ian Connick Covert, Wei Qiu, Mingyu Lu, Na Yoon Kim, Nathan J. White, and Su-In Lee. Learning to Maximize Mutual Information for Dynamic Feature Selection. *ICML*, PMLR 202:6424–6447, 2023. https://proceedings.mlr.press/v202/covert23a.html.

A. Deshpande, L. Hellerstein, and D. Kletenik. Approximation Algorithms for Stochastic Boolean Function Evaluation and Stochastic Submodular Set Cover. *SODA*, pp. 1453–1466, 2014. https://doi.org/10.1137/1.9781611973402.107.

Aurélien Garivier and Emilie Kaufmann. Optimal Best Arm Identification with Fixed Confidence. *COLT*, PMLR 49:998–1027, 2016. https://proceedings.mlr.press/v49/garivier16a.html.

K. Gatmiry, T. Kesselheim, S. Singla, and Y. Wang. Bandit Algorithms for Prophet Inequality and Pandora's Box. *SODA*, pp. 462–500, 2024. https://doi.org/10.1137/1.9781611977912.18.

R. Ghuge, A. Gupta, and V. Nagarajan. Nonadaptive Stochastic Score Classification and Explainable Half-Space Evaluation. *Operations Research*, 73(4):2204–2222, 2025; first published online 18 July 2024. Preliminary version in IPCO 2022. https://doi.org/10.1287/opre.2023.0431.

R. Ghuge, J. Moondra, and M. Singh. Stochastic Function Certification with Correlations. arXiv:2604.02611, 2026. Preprint. https://arxiv.org/abs/2604.02611.

D. Gkenosis, N. Grammel, L. Hellerstein, and D. Kletenik. The Stochastic Score Classification Problem. *ESA*, LIPIcs 112:36:1–36:14, 2018. https://doi.org/10.4230/LIPIcs.ESA.2018.36.

D. Gkenosis, N. Grammel, L. Hellerstein, and D. Kletenik. The Stochastic Boolean Function Evaluation Problem for Symmetric Boolean Functions. *Discrete Applied Mathematics*, 309:269–277, 2022. https://doi.org/10.1016/j.dam.2021.12.001.

Osman Berke Guney, Ketan Suhaas Saichandran, Karim Elzokm, Ziming Zhang, and Vijaya B. Kolachalama. Active feature acquisition via explainability-driven ranking. *ICML*, PMLR 267:20748–20765, 2025. https://proceedings.mlr.press/v267/guney25a.html.

C. Guo, Z. Huang, Z. G. Tang, and X. Zhang. Generalizing Complex Hypotheses on Product Distributions: Auctions, Prophet Inequalities, and Pandora's Problem. *COLT*, PMLR 134:2248–2288, 2021. https://proceedings.mlr.press/v134/guo21a.html.

A. Gupta, H. Jiang, Z. Scully, and S. Singla. The Markovian Price of Information. *IPCO*, LNCS 11480:233–246, 2019. https://doi.org/10.1007/978-3-030-17953-3_18.

C. Jin, T. Jin, H. Luo, S. Sra, and T. Yu. Learning Adversarial Markov Decision Processes with Bandit Feedback and Unknown Transition. *ICML*, PMLR 119:4860–4869, 2020. https://proceedings.mlr.press/v119/jin20c.html.

T. Juškevičius. The sharp form of the Kolmogorov–Rogozin inequality and a conjecture of Leader–Radcliffe. *Bulletin of the London Mathematical Society*, 56(11):3289–3299, 2024. https://doi.org/10.1112/blms.13132.

J. Kossen, S. Farquhar, Y. Gal, and T. Rainforth. Active Testing: Sample-Efficient Model Evaluation. *ICML*, PMLR 139:5753–5763, 2021. https://proceedings.mlr.press/v139/kossen21a.html.

Yang Li, Jie Ma, Miguel Ballesteros, Yassine Benajiba, and Graham Horwood. Active Evaluation Acquisition for Efficient LLM Benchmarking. *ICML*, PMLR 267:35581–35602, 2025. https://proceedings.mlr.press/v267/li25bp.html.

Jiawei Liu, Chunqiu Steven Xia, Yuyao Wang, and Lingming Zhang. Is Your Code Generated by ChatGPT Really Correct? Rigorous Evaluation of Large Language Models for Code Generation. *Advances in Neural Information Processing Systems*, 36, 2023. https://papers.neurips.cc/paper_files/paper/2023/hash/43e9d647ccd3e4b7b5baab53f0368686-Abstract-Conference.html.

B. M. Plank and K. Schewior. Simple Algorithms for Stochastic Score Classification with Small Approximation Ratios. *SIAM Journal on Discrete Mathematics*, 38(3):2069–2088, 2024. https://doi.org/10.1137/22M1523492.

N. Ross. Fundamentals of Stein's method. *Probability Surveys*, 8:210–293, 2011. https://doi.org/10.1214/11-PS182.

S. Salloum and M. A. Breuer. An optimum testing algorithm for some symmetric coherent systems. *Journal of Mathematical Analysis and Applications*, 101(1):170–194, 1984. https://doi.org/10.1016/0022-247X(84)90064-7.

TensorFlow Authors. The Evaluator TFX Pipeline Component. Official documentation, updated September 6, 2024. https://www.tensorflow.org/tfx/guide/evaluator.

T. Ünlüyurt. Sequential testing of complex systems: a review. *Discrete Applied Mathematics*, 142(1–3):189–205, 2004. https://doi.org/10.1016/j.dam.2002.08.001.

A. Vera and S. Banerjee. The Bayesian Prophet: A Low-Regret Framework for Online Decision Making. *Management Science*, 67(3):1368–1391, 2021. https://doi.org/10.1287/mnsc.2020.3624.

M. L. Weitzman. Optimal Search for the Best Alternative. *Econometrica*, 47(3):641–654, 1979. https://doi.org/10.2307/1910412.

Terry Yue Zhuo et al. BigCodeBench: Benchmarking Code Generation with Diverse Function Calls and Complex Instructions. *ICLR*, 2025. https://proceedings.iclr.cc/paper_files/paper/2025/hash/a6a90bcc2aa470c3871b2d39a67d26e8-Abstract-Conference.html.

A. Zimin and G. Neu. Online learning in episodic Markovian decision processes by relative entropy policy search. *NeurIPS*, 26, 2013. https://proceedings.neurips.cc/paper/2013/hash/68053af2923e00204c3ca7c6a3150cf7-Abstract.html.

Xia Zong, Tuomo Lehtonen, and Jussi Rintanen. Cost-Optimal Decision Diagrams for Stochastic Boolean Function Evaluation. *arXiv:2606.24672*, 2026. Preprint. https://arxiv.org/abs/2606.24672.

APPENDIX CONTENTS

The appendices follow the proof dependencies. Appendix A establishes the observation model. Appendices B and C prove the cost comparison. Appendix D proves the learning bound and its lower rate. Appendix E gives the finite calculations and protocols.

⋆ Original results and constructions; † adaptations or refinements of existing methods; ○ standard methods; ◇ models, examples and experimental protocols. Section marks give the main role; subsection marks identify the ingredients.

† **A Observation model and stopped records** 13
† A.1 Stopped records and conditional product laws 14
⋆ **B Local prices and the cost lower bound** 15
◇ B.1 Local response model 15
† B.2 Net columns and their price envelope 16
◇ B.3 Count evidence from required action queries 17
◇ B.4 A fixed code-evaluation interface 18
○ B.5 A catalog generating the local prices 19
† B.6 Boundary exposure and cost balance 20
⋆ B.7 Component-excluding prices and the lower bound 21
⋆ **C Executing the allocation and proving the cost comparison** 26
† C.1 Reserve selection and exact completion 26
⋆ C.2 Conditional delayed-activation tracking 28
⋆ C.3 Library supply and the count scale 35
⋆ C.4 Proof of the cost comparison and supply examples 37
⋆ **D Learning on the original batch** 39
† D.1 Executed feedback and exact certification 39
† D.2 Learning on the query-history graph 41
† D.3 Expected regret of the graph learner 42
⋆ D.4 Reserve selection and the learning upper bound 45
⋆ D.5 Local costs and information in the two-sensor family 49
⋆ D.6 The fourth-root lower bound and matching upper rate 52
◇ **E Finite query costs and recorded evidence** 54
○ E.1 Exact finite costs in the one-hint interface 54
◇ E.2 Learning fees and fixed-record protocols 55

The original results are the cost comparison over all legal queries, its delayed-activation implementation, and the matching learning-cost bounds. The supporting adaptations connect occupation measures, stopped-cost accounting, online price updates and confidence bounds to this query model. Bellman recursion, linear-programming duality, concentration inequalities and zero-bias normal approximation are standard tools used in these arguments.

NOTATION

Throughout the appendices, a *root* means a physical component. A fresh or untouched root has not yet been queried and retains its original response law. The cost comparison and learning bound use the same endpoint correction.

Symbol	Meaning throughout the core manuscript
N, n	Physical components; maximal total integer contribution
A_g, X_g	Required local action; selected count contribution
K, Γ	Total contribution; fixed category map
$\mathfrak{h}_g, \mathsf{R}_T$	Local history; complete stopped query record
$\mathscr{F}_T, \mathscr{G}_{g,T}$	Stopped filtration; enlargement by other full mechanisms
C_π, v_g, V_0, C_N^*	Paid fee; full local value; baseline; joint optimum
$B_\pi^\pm, G_\pi$	Net directional exposures and gain
$\lambda, \widehat{\rho}_\pm$	Dual prices; axial net prices
$\mathcal{H}_N, \mathcal{V}, \mathcal{F}^{\mathrm{net}}$	Actual supply value; cone value; count-dependent value
$\mathcal{C}, \beta_{\mathrm{gen}}$	Chosen generating catalog and its coverage
$D_k^\pm, D_N$	Expected count margins, not query records
θ, d	Shared quality parameter; local two-law separation
$\chi_{\mathrm{rec}}, \beta_0$	Supplied completion price and covered-policy exposure
$\mathfrak{D}_N, \zeta_N, \tau_N$	Library shortfall, count geometry, intrinsic reserve scale
c_{gate}	Fixed bound on the normalized count-to-threshold distance
$c_{\mathrm{low}}, c_{\mathrm{up}}, c_{\mathrm{learn}}$	Named cost-bound constants, distinct from the count K

A OBSERVATION MODEL AND STOPPED RECORDS

We specify which mechanisms remain possible after a query record. This gives the stopping criterion and a finite cost certificate. For independent components, the final subsection proves the conditional product law used throughout the cost and learning arguments.

Responses and admissible outputs. A finite mechanism $\omega \in \Omega$ fixes the address responses $X_e(\omega) \in \mathcal{X}_e$, $e \in E$. A transcript is $\mathsf{R} = (Q, x_Q)$; $\mathcal{Q}(\mathsf{R}) \subseteq E \setminus Q$ specifies currently legal queries. A response is cached and charged once:

$$\Omega_{\mathsf{R}} = \{\omega : X_Q(\omega) = x_Q\}, \qquad w(Q) = \sum_{e \in Q} w_e, \quad w_e > 0, \tag{16}$$

$$\mathcal{Z}_{\mathsf{R}} = \bigcap_{\omega \in \Omega_{\mathsf{R}}} \mathcal{Z}(\omega), \qquad \mathsf{T} = \{\mathsf{R} : \mathcal{Z}_{\mathsf{R}} \neq \varnothing\}. \tag{17}$$

For a fixed local action target and aggregate, the joint target is

$$\mathcal{Z}(\omega) = \{((A_g(\omega_g))_{g=1}^N, \Gamma(\sum_g X_g(\omega_g)))\}.$$

A shared address is one element of E. Observations are passive; a fixed legal full evaluation identifies the requested output. For expected cost, a prior $\mu(\omega) > 0$ on the declared support gives

$$\mu_{\mathsf{R}}(\omega) = \frac{\mu(\omega)\mathbf{1}\{\omega \in \Omega_{\mathsf{R}}\}}{\mu(\Omega_{\mathsf{R}})}, \qquad p_{\mathsf{R}}(e, x) = \mu_{\mathsf{R}}\{X_e = x\}. \tag{18}$$

Correctness is required for every compatible mechanism. The prior governs query selection and expected fees.

Product specialization. For independent components, $E = \bigsqcup_g E_g$ and $\Omega_{\mathsf{R}} = \prod_g \Omega_{g,\mathsf{R}}$. Unresolved binary losses then have attainable count endpoints $[S, S + R]$, where S counts established ones and R unresolved losses. The count classifier has distinct integer cells $[a, b]$. Adaptive revelation leaves the sampled batch fixed; exactness refers to its prescribed output.

Finite optimality certificate. The standard Bellman inequalities on the finite legal history graph are

$$\begin{aligned} L(\mathsf{R}) &= 0 && (\mathsf{R} \in \mathsf{T}), \\ L(\mathsf{R}) &\le w_e + \sum_x p_{\mathsf{R}}(e,x) L(\mathsf{R} \cup \{e = x\}) && (\mathsf{R} \notin \mathsf{T},\ e \in \mathcal{Q}(\mathsf{R})). \end{aligned} \tag{19}$$

Let $\tau \le |E|$ be the stopping time of an exact policy. Put $D_t = w_{e_t} + \mathbb{E}[L(\mathsf{R}_{t+1}) \mid \mathscr{F}_t] - L(\mathsf{R}_t)$ on $\{t < \tau\}$. Then $D_t \ge 0$ by (19), and

$$\begin{aligned} \mathbb{E}_\mu \sum_{t<\tau} D_t &= \mathbb{E}_\mu C_\pi + \sum_{t=0}^{|E|-1} \mathbb{E}_\mu\big[\mathbf{1}_{\{t<\tau\}} \mathbb{E}_\mu[L(\mathsf{R}_{t+1}) - L(\mathsf{R}_t) \mid \mathscr{F}_t]\big] \\ &\overset{\text{tower}}{=} \mathbb{E}_\mu C_\pi + \mathbb{E}_\mu \sum_{t<\tau} [L(\mathsf{R}_{t+1}) - L(\mathsf{R}_t)] \\ &= \mathbb{E}_\mu C_\pi + \underbrace{\mathbb{E}_\mu L(\mathsf{R}_\tau)}_{0 \text{ by exact stopping}} - L(\mathsf{R}_0) \ge 0. \end{aligned} \tag{20}$$

Backward minimization chooses $D_t = 0$ at each queried node. For that policy $\mathbb{E}_\mu C_\pi = L(\mathsf{R}_0)$.

Finite observable patterns. A latent object z_g need not have finite support. If the query addresses have finite response alphabets and the requested local targets have finite ranges, map it to

$$T_g(z_g) = ((X_e(z_g))_{e \in E_g}, A_g(z_g), X_g(z_g)).$$

The image is finite. Its law is the push-forward of the latent law. For a fixed private seed, the record identity propagates as

$$\mathsf{R}_t(z) = \mathsf{R}_t(T(z)) \implies e_{t+1}(z) = e_{t+1}(T(z)) \implies X_{e_{t+1}}(z) = X_{e_{t+1}}(T(z)).$$

Starting at the empty record gives identical costs and stops. For public context W, conditional use requires $\mu(d\omega \mid W) = \bigotimes_g \mu_g(d\omega_g \mid W)$. A conditional bound $\mathbb{E}[C \mid W] \le B(W)$ gives $\mathbb{E}C = \mathbb{E}[\mathbb{E}(C \mid W)] \le \mathbb{E}B(W)$.

A.1 Stopped records and conditional product laws

Let U contain all private randomization and be independent of $\omega = (\omega_1, \dots, \omega_N)$. A randomized policy is deterministic conditional on U. At time t, its next queried address is measurable in

$$\mathscr{F}_t = \sigma(U, e_1, R_1, \dots, e_t, R_t), \qquad R_j = X_{e_j}(\omega),$$

and its stopping time T is bounded by the number of distinct addresses. After T pad the record by a fixed null symbol. For a possible finite stopped record $d = (e_1, r_1, \dots, e_t, r_t, T = t)$, let $\Omega_g(d) = \{\omega_g : X_{e_j}(\omega_g) = r_j \text{ for all } e_j \in E_g\}$. For every fixed seed u, there is a deterministic $a(u,d) \in \{0,1\}$ such that

$$\mathbf{1}_{\{\mathsf{R}_T = d\}}(u,\omega) = a(u,d) \prod_{g=1}^{N} \mathbf{1}_{\Omega_g(d)}(\omega_g). \tag{21}$$

Indeed the first address, each subsequent choice and the decision to stop are fixed by u and the preceding entries of d. If any of these choices disagrees with d, set $a(u,d) = 0$. Otherwise a mechanism generates d exactly when it returns all listed responses, which is the displayed product. For a record d and a measurable seed event B, put $a_B(d) = \int_B a(u,d) P_U(du)$. Independence gives

$$\begin{aligned} \Pr(\omega = w, \mathsf{R}_T = d, U \in B) &= a_B(d) \prod_g \mu_g(w_g) \mathbf{1}_{\Omega_g(d)}(w_g), \\ \Pr(\mathsf{R}_T = d, U \in B) &= a_B(d) \prod_g \mu_g(\Omega_g(d)). \end{aligned}$$

Taking the ratio on positive-probability records gives

$$\Pr(\omega \in d\boldsymbol{w} \mid \mathscr{F}_T) = \bigotimes_{g=1}^{N} \frac{\mathbf{1}_{\Omega_g(\mathsf{R}_T)}(w_g) \mu_g(dw_g)}{\mu_g(\Omega_g(\mathsf{R}_T))}. \tag{22}$$

The factor $a_B(d)$ cancels. Taking the whole seed space gives the formula without the seed. Define

$$m_g = \mathbb{E}[X_g \mid \mathscr{F}_T], \qquad \mathscr{G}_{g,T} = \mathscr{F}_T \vee \sigma(\omega_j : j \neq g).$$

For a bounded $\mathscr{G}_{g,T}$-measurable Z_g,

$$\begin{aligned} \Pr(\omega_g = w \mid \mathscr{G}_{g,T}) &= \frac{\mathbf{1}_{\Omega_g(\mathsf{R}_T)}(w)\mu_g(w)}{\mu_g(\Omega_g(\mathsf{R}_T))}, \\ \mathbb{E}[X_g \mid \mathscr{G}_{g,T}] &= \sum_w X_g(w)\frac{\mathbf{1}_{\Omega_g(\mathsf{R}_T)}(w)\mu_g(w)}{\mu_g(\Omega_g(\mathsf{R}_T))} = m_g, \\ \mathbb{E}[(m_g - X_g)Z_g] &\overset{\text{tower}}{=} \mathbb{E}[Z_g\{m_g - \mathbb{E}[X_g \mid \mathscr{G}_{g,T}]\}] = 0. \end{aligned} \tag{23}$$

For an unqueried component $\Omega_g(\mathsf{R}_T) = \Omega_g$. Its conditional law is therefore its original law. These statements condition on the complete record.

B Local prices and the cost lower bound

We express each local policy by its paid fee and unresolved count exposure. The resulting price envelope supplies a lower bound on the cost of every exact batch policy. The action-query and code-evaluation examples show how to obtain these quantities from a specified interface.

B.1 Local response model

Use the action targets and aggregate outcomes defined in Section 3. The common aggregate bound is $x_{\max}$. **(P)** Components have disjoint addresses and independent full-support laws:

$$E = \bigsqcup_{g=1}^{N} E_g, \qquad \Omega = \prod_g \Omega_g, \qquad \mu = \bigotimes_g \mu_g, \qquad \mu_g(\omega_g) > 0. \tag{24}$$

Inside a component, responses and targets may be arbitrarily related. The finite local table specifies responses, targets, costs, the prior and legal queries $Q_g(\mathfrak{h})$, not the realized mechanism. Local access prerequisites are allowed. Every reachable history admits a legal completion identifying (A_g, X_g); in bounded-fee learning classes its pathwise fee is at most C_{loc}. The stopped-record factorization in Appendix A.1 gives

$$\begin{aligned} \mu(\omega \mid \mathfrak{h}) = \prod_g \mu_g(\omega_g \mid \mathfrak{h}_g), \quad m_g(\mathfrak{h}) = \mathbb{E}[X_g \mid \mathfrak{h}], \\ l_g(\mathfrak{h}) = \min X_g(\Omega_{g,\mathfrak{h}}), \qquad u_g(\mathfrak{h}) = \max X_g(\Omega_{g,\mathfrak{h}}). \end{aligned} \tag{25}$$

Let $\mathcal{Z}_g$ identify (A_g, X_g) and $\mathcal{S}_g$ identify A_g but not X_g. Only these histories occur at joint termination. On $\mathcal{S}_g$, $b_g^+ = m_g - l_g > 0$, $b_g^- = u_g - m_g > 0$; both vanish on $\mathcal{Z}_g$. The local full-evaluation value and query excess cost are

$$\begin{aligned} v_g(\mathfrak{h}) &= 0 && (\mathfrak{h} \in \mathcal{Z}_g), \\ v_g(\mathfrak{h}) &= \min_{e \in Q_g(\mathfrak{h})} \{w_e + \sum_x p_g(x \mid \mathfrak{h}, e) v_g(\mathfrak{h}^{e,x})\} && (\mathfrak{h} \notin \mathcal{Z}_g), \end{aligned} \tag{26}$$

$$d_g(\mathfrak{h}, e) = w_e + \mathbb{E}[v_g(\mathfrak{h}^{e,x}) \mid \mathfrak{h}, e] - v_g(\mathfrak{h}) \geq 0. \tag{27}$$

For d binary addresses there are at most 3^d local histories. Define the directional completion prices

$$\mathfrak{c}_\pm = \max_{g, \mathfrak{h} \in \mathcal{S}_g} \frac{v_g(\mathfrak{h})}{b_g^\pm(\mathfrak{h})}. \tag{28}$$

Empty maxima are zero; empty eligible catalogs give exact cost V_0.

Call a component *endpoint-preserving* when every known-action, unresolved history retains its original extreme attainable outcomes:

$$l_g(\mathfrak{h}) = l_g(\mathfrak{h}_g^0), \qquad u_g(\mathfrak{h}) = u_g(\mathfrak{h}_g^0). \tag{EP}$$

Binary aggregates satisfy (EP). So do multivalued outcomes observed through strictly positive noisy channels until a direct outcome query resolves them; posterior masses may approach zero without removing either endpoint.

B.2 NET COLUMNS AND THEIR PRICE ENVELOPE

Let Π_j be the finite deterministic local partial policies starting at type j's root and stopping only with known actions. They may leave X_j unknown. Each policy is tagged by its type, including that type's law and fees, as in Section 3. For stopping history $\mathfrak{h}_\pi$ and paid cost C_π, define

$$B_\pi^{\pm} = \mathbb{E}b_j^{\pm}(\mathfrak{h}_\pi), \qquad G_\pi = v_j(\mathfrak{h}_j^0) - \mathbb{E}C_\pi = \mathbb{E}v_j(\mathfrak{h}_\pi) - \mathbb{E}\sum_{t<\tau_\pi} d_j(\mathfrak{h}_t, e_t). \tag{29}$$

For the last equality, write the stopped sum explicitly:

$$\begin{aligned}\mathbb{E}\sum_{t<\tau_\pi} d_j(\mathfrak{h}_t, e_t) &= \mathbb{E}\sum_{t<\tau_\pi} w_{e_t} + \sum_t \mathbb{E}\big[\mathbf{1}_{\{t<\tau_\pi\}}\mathbb{E}[v_j(\mathfrak{h}_{t+1}) - v_j(\mathfrak{h}_t) \mid \mathscr{F}_t]\big] \\ &= \mathbb{E}C_\pi + \mathbb{E}\sum_{t<\tau_\pi}[v_j(\mathfrak{h}_{t+1}) - v_j(\mathfrak{h}_t)] \\ &= \mathbb{E}C_\pi + \mathbb{E}v_j(\mathfrak{h}_\pi) - v_j(\mathfrak{h}_j^0).\end{aligned}$$

Every paid branch enters this identity. Randomization gives convex mixtures. Let $\Pi_+ = \{\pi : G_\pi > 0\}$. Full support gives

$$\begin{aligned}B_\pi^s = 0 &\implies b_j^s(\mathfrak{h}_\pi) = 0 \text{ a.s.} \implies \mathfrak{h}_\pi \in \mathcal{Z}_j \text{ a.s.} \\ &\implies \mathbb{E}C_\pi \geq v_j(\mathfrak{h}_j^0) \implies G_\pi \leq 0.\end{aligned}$$

Hence every column in Π_+ has both exposures positive.

For nonnegative capacities, with absent sides omitted, set

$$\mathcal{V}(c_+, c_-) = \max_{z_\pi \geq 0}\left\{\sum_{\pi\in\Pi_+} G_\pi z_\pi : \sum_\pi B_\pi^+ z_\pi \leq c_+, \quad \sum_\pi B_\pi^- z_\pi \leq c_-\right\}. \tag{30}$$

At x in the count cell $[a, b]$, write $\mathcal{F}^{\mathrm{net}}(x) = \mathcal{V}(x - a, b - x)$, omitting $a = 0$ or $b = n$; set it to zero in the open unit gaps between cells. Its dual and densities are

$$\begin{aligned}\mathcal{D} &= \{\lambda \geq 0 : \lambda_+ B_\pi^+ + \lambda_- B_\pi^- \geq G_\pi \quad \forall \pi \in \Pi_+\}, \\ \mathcal{V}(c_+, c_-) &= \min_{\lambda\in\mathcal{D}}(\lambda_+ c_+ + \lambda_- c_-), \qquad (31)\end{aligned}$$

$$\widehat{\rho}_\pm = \max_{\pi\in\Pi_+} G_\pi / B_\pi^{\pm}, \qquad L_{\mathrm{net}} = \widehat{\rho}_+ + \widehat{\rho}_-. \tag{32}$$

An absent side sets its multiplier to zero. Clipping prices to $[0, \widehat{\rho}_+] \times [0, \widehat{\rho}_-]$ preserves feasibility, so $\mathcal{F}^{\mathrm{net}}$ is L_{net}-Lipschitz, including its zero-valued internal endpoints. A basic optimum uses at most two policies. In particular,

$$\mathcal{F}^{\mathrm{net}}(K) = \widehat{\rho}_+(K - k)_+ + \widehat{\rho}_-(k - 1 - K)_+ \quad \text{for one threshold.} \tag{33}$$

Computing the envelope on the history graph. Let t_j be type j's root flow, $x_\mathfrak{h}$ its arriving flow, $y_{\mathfrak{h}e}$ query flow, and $s_\mathfrak{h}$ stopping flow. On the finite local history graph impose

$$\begin{aligned}x_\mathfrak{h} &= t_j \mathbf{1}\{\mathfrak{h} = \mathfrak{h}_j^0\} + \sum_{\widetilde{\mathfrak{h}}, e, z : \widetilde{\mathfrak{h}}^{e,z} = \mathfrak{h}} p_j(z \mid \widetilde{\mathfrak{h}}, e) y_{\widetilde{\mathfrak{h}}e}, \\ x_\mathfrak{h} &= s_\mathfrak{h} + \sum_{e\in Q_j(\mathfrak{h})} y_{\mathfrak{h}e}, \qquad s_\mathfrak{h} = 0 \quad (\mathfrak{h} \notin \mathcal{Z}_j \cup \mathcal{S}_j),\end{aligned} \tag{34}$$

with nonnegative variables and no outgoing flow from resolved states. Maximizing $\sum_j t_j v_j(\mathfrak{h}_j^0) - \sum_{j,\mathfrak{h},e} w_e y_{\mathfrak{h}e}$ under $\sum_{j,\mathfrak{h}} s_\mathfrak{h} b_j^{\pm}(\mathfrak{h}) \leq c_\pm$ equals (30). For $t_j > 0$, stop with probability $s_\mathfrak{h}/x_\mathfrak{h}$ or query e with probability $y_{\mathfrak{h}e}/x_\mathfrak{h}$. Zero-flow nodes are never visited. At the root, $\Pr_\pi(\mathfrak{h}_j^0 \text{ visited}) = 1 = x_{\mathfrak{h}_j^0}/t_j$. If the visitation identity holds at every predecessor of $\mathfrak{h}$, then

$$\begin{aligned}\Pr_\pi(\mathfrak{h} \text{ visited}) &= \sum_{\widetilde{\mathfrak{h}}, e, z : \widetilde{\mathfrak{h}}^{e,z} = \mathfrak{h}} \frac{x_{\widetilde{\mathfrak{h}}}}{t_j}\frac{y_{\widetilde{\mathfrak{h}}e}}{x_{\widetilde{\mathfrak{h}}}} p_j(z \mid \widetilde{\mathfrak{h}}, e) \\ &= \frac{1}{t_j}\sum_{\widetilde{\mathfrak{h}}, e, z} p_j(z \mid \widetilde{\mathfrak{h}}, e) y_{\widetilde{\mathfrak{h}}e} = x_\mathfrak{h}/t_j.\end{aligned}$$

Thus

$$\Pr_\pi(\mathfrak{h} \text{ visited}) = x_\mathfrak{h}/t_j,$$
$$\Pr_\pi(\mathfrak{h}, e \text{ queried}) = y_{\mathfrak{h}e}/t_j, \qquad \Pr_\pi(\mathfrak{h}_\pi = \mathfrak{h}) = s_\mathfrak{h}/t_j,$$
$$t_j \mathbb{E}C_\pi = \sum_{\mathfrak{h},e} w_e y_{\mathfrak{h}e}, \qquad t_j B^s_\pi = \sum_{\mathfrak{h}} s_\mathfrak{h} b^s_j(\mathfrak{h}).$$

Conversely these visitation probabilities satisfy (34). This is the finite occupation-measure representation (Altman, 1996).

A dual price may instead be checked by local priced stopping:

$$U_{j,\lambda}(\mathfrak{h}) = \min\left\{\lambda \cdot b_j(\mathfrak{h}) \text{ if actions are known}, \quad \min_{e \in Q_j(\mathfrak{h})} [w_e + \mathbb{E}U_{j,\lambda}(\mathfrak{h}^{e,X})]\right\}. \tag{35}$$

At resolved states $U = 0$. At the root,

$$\begin{aligned} U_{j,\lambda}(\mathfrak{h}^0_j) &= \min_{\pi \in \Pi_j} \{\mathbb{E}C_\pi + \lambda \cdot B_\pi\} \\ &= v_j(\mathfrak{h}^0_j) - \max_{\pi \in \Pi_j} \{G_\pi - \lambda \cdot B_\pi\}. \end{aligned}$$

Full evaluation has $(B, G) = (0, 0)$. Thus the maximum is nonnegative and $U = v$ precisely when every $G_\pi \le \lambda \cdot B_\pi$. Testing all types is equivalent to $\lambda \in \mathcal{D}$.

Global count conditions and local controller parameters. The global count condition used below is

$$\begin{aligned} \text{(K)} \qquad & X_g \in \{0, \dots, x_{\max}\}, \quad \mu_N = \mathbb{E}K, \quad \sigma^2_N = \operatorname{Var} K \ge \nu N, \\ & 1 \le q_\Gamma \le q_0, \qquad \max_{1 \le s \le q_\Gamma} |k_{s,N} - \mu_N| \le c_{\text{gate}} \sqrt{N}. \end{aligned} \tag{36}$$

Here $x_{\max}, \nu, c_{\text{gate}}, q_0$ are fixed, and q_Γ is the number of genuine count boundaries. No type frequency or local graph-size bound is part of (K). A tracking subproblem instead specifies a fixed finite menu of J local types and an actual number of untouched roots per type. All its constants refer only to this menu. Lemma C treats the conditional subproblem; Lemma D supplies it from the batch. The learning class and its coverage constants are defined separately in Appendix D.4.

B.3 Count evidence from required action queries

Consider the fair-bit hint interface of Section 4, now with the initially unknown required action $A = H$. Write $\varepsilon = \varepsilon_h$, with $0 < \varepsilon \le 1/2$ fixed. The mandatory action fee remains in the baseline, giving

$$(B^+, B^-, G) = (\varepsilon/2, (1-\varepsilon)/2, r/2), \qquad \widehat{\rho}_+ = \widehat{\rho}_- = r/\varepsilon. \tag{37}$$

Its four positive masses are

$$\Pr(Y = y, H = h) = \tfrac{1}{2}\big[(1-\varepsilon)\mathbf{1}_{\{y=h\}} + \varepsilon \mathbf{1}_{\{y \ne h\}}\big].$$

Since $\ell(a) = Y + \gamma \mathbf{1}_{\{a \ne H\}}$, $\gamma > 0$, one has $\arg\min_{a \in \{0,1\}} \ell(a) = H$ and $\min_a \ell(a) = Y$. On every history omitting the H response, both action values remain compatible. On every history omitting the Y response, both selected losses remain compatible. Thus full evaluation reads both addresses on every mechanism and $v(\mathfrak{h}^0) = c_h + r$.

A deterministic action-sufficient policy that starts with Y must subsequently read H; its net column is zero. A policy starting with H is determined by a set $R \subseteq \{0, 1\}$ of hint values on which it retains Y. Set $p_0 = \varepsilon$ and $p_1 = 1 - \varepsilon$. Its fee and column are

$$\begin{aligned} \mathbb{E}C_R &= c_h + r(1 - |R|/2), \quad G_R = r|R|/2, \\ B^+_R &= \tfrac{1}{2}\sum_{h \in R} p_h, \qquad\qquad B^-_R = \tfrac{1}{2}\sum_{h \in R}(1 - p_h). \end{aligned} \tag{38}$$

These four sets exhaust deterministic policies after deleting redundant post-completion operations. Randomization gives their convex hull. Hence

$$\begin{aligned} R = \{0\} &: (B^+, B^-, G) = (\varepsilon/2, (1-\varepsilon)/2, r/2), \\ R = \{1\} &: (B^+, B^-, G) = ((1-\varepsilon)/2, \varepsilon/2, r/2), \\ R = \{0,1\} &: (B^+, B^-, G) = (1/2, 1/2, r). \end{aligned}$$

The third column is the sum of the first two. For either coordinate, the ratio of a nonnegative sum of positive columns is a weighted average of their individual ratios, with weights proportional to that coordinate. Since $r/\varepsilon \geq r/(1-\varepsilon)$, this proves the axial prices $\widehat{\rho}_+ = \widehat{\rho}_- = r/\varepsilon$ over the entire query class. Both positive selective columns have exposure at least $\varepsilon/2$, and the binary support condition holds. For odd N and majority $k = (N+1)/2$, put $D_N = \mathbb{E}|K - N/2| - 1/2$. With N independent copies, The comparison has $V_0 = N(c_h + r)$, $\mathfrak{D}_N = 0$, $\zeta_N = 0$ and sufficient supply $m_0 = N$ for all large N. Hence

$$\begin{aligned} \mathbb{E}F_0(K) &= (r/\varepsilon)\mathbb{E}[(K-k)_+ + (k-1-K)_+] = (r/\varepsilon)D_N, \\ N(c_h + r) - (r/\varepsilon)D_N \leq C_N^*(\varepsilon) &\leq N(c_h + r) - (r/\varepsilon)D_N + c_{\mathrm{up}} + 2\Lambda_0. \end{aligned}$$

Writing the bounded remainder as $R_N(\varepsilon)$ gives

$$\begin{aligned} C_N^*(\varepsilon) &= N(c_h + r) - (r/\varepsilon)D_N + R_N(\varepsilon), \\ C_N^*(1/2) - C_N^*(\varepsilon) &= (r/\varepsilon - 2r)D_N + R_N(1/2) - R_N(\varepsilon) \\ &= (r/\varepsilon - 2r)D_N + O(1). \end{aligned} \tag{39}$$

All parameters, including ε, are fixed in these expansions. For the learning class with $\varepsilon \in [\varepsilon_{\min}, 1/2]$, $0 < \varepsilon_{\min} \leq 1/2$, every unresolved action-sufficient history satisfies $\min(b^+, b^-) \geq \varepsilon_{\min}$, and the two selective columns satisfy $\min(B^+, B^-) \geq \varepsilon_{\min}/2$. They generate the whole dual region by (38); the support graph and minimizing full routine are common to the class. These are precisely the coverage and interface premises of Theorem 2. The baseline includes the mandatory action query in both laws. The optional hint example instead has baseline r and selective gain $r/2 - c_h$; it is therefore a different information contract even when the sensor is the same.

For that optional-hint interface, a deterministic policy either stops before either query, reads the result, or first reads the sole hint. In the third case it retains a set $R \subseteq \{0,1\}$ of hint responses. Its exposures are as in (38), with gain $r|R|/2 - c_h$. The cases $R = \varnothing$ and $R = \{0,1\}$ are dominated by a direct result query and immediate retention, respectively. The two selective policies give the ratios $(r - 2c_h)/\varepsilon_{\mathrm{h}}$ and $(r - 2c_h)/(1 - \varepsilon_{\mathrm{h}})$. For $0 < \varepsilon_{\mathrm{h}} \leq 1/2$, taking their maximum together with the immediate-retention ratio $2r$ proves (5).

B.4 A FIXED CODE-EVALUATION INTERFACE

Freeze a generated program, a test version, execution seeds, timeout rules and the proxy model and its seed for each item. Let H be the proxy response, P the pass bit for a fixed test prefix, and X the pass bit for the complete suite. The aggregate is $K = \sum_g X_g$. A release bit is a specified threshold function of K; requiring that bit does not require exporting every X_g. The support is the six-element set

$$\Omega_{\mathrm{code}} = \{(h, p, x) \in \{0,1\}^3 : x \leq p\}.$$

The scheduling policy receives immutable item identifiers, the protocol and paid responses. Source inspection and individual test results, when exposed by a system, are additional legal observations whose fees and support refinements belong in its query graph. The present graph uses prefix and remainder jobs as query units. The available queries are: obtain H at fee c_h; run the prefix to obtain P at fee c_p; after $P = 1$, run the remainder to obtain X at fee r. When $P = 0$, $X = 0$ is already certified. At an unresolved history the proxy may be obtained before or after the prefix. A query is cached. Local actions are fixed in this example; the joint-action family treats unknown actions. The prefix and remainder are atomic jobs of this specified interface. A system exposing individual test results is a finer graph and is analyzed by replacing these jobs with its actual legal queries.

If every atom has positive mass, the proxy never resolves X. Any exact local completion must execute the prefix and, on its passing branch, the remainder. For any exact full policy σ,

$$\begin{aligned} C_\sigma &\geq c_p + r\mathbf{1}_{\{P=1\}}, \\ \inf_\sigma \mathbb{E}C_\sigma &\geq c_p + r\Pr(P=1) = \mathbb{E}C_{\text{prefix then remainder}} = v(\mathfrak{h}^0). \end{aligned} \tag{40}$$

The displayed routine attains the lower bound under every positive law. This supplies a common optimal full routine with unknown expected cost. The support after either hint is $[0,1]$, after prefix failure is $[0,0]$, and after the remainder is a singleton. This verifies the support-anchor condition. Unknown response frequencies affect efficiency, whereas support containment determines whether stopping is legal.

For example, first read H, retain if $H = 1$, and otherwise use the full routine. Its terminal interval is $[0, 1]$ on $\{H = 1\}$ and $\{X\}$ otherwise. Consequently

$$\begin{aligned} B^+ &= \mathbb{E}[(X-0)\mathbf{1}_{\{H=1\}}] = \Pr(H=1, X=1), \\ B^- &= \mathbb{E}[(1-X)\mathbf{1}_{\{H=1\}}] = \Pr(H=1, X=0), \\ \mathbb{E}C_\pi &= c_h + c_p \Pr(H=0) + r \Pr(H=0, P=1), \\ G &= c_p + r \Pr(P=1) - \mathbb{E}C_\pi \\ &= c_p \Pr(H=1) + r \Pr(H=1, P=1) - c_h. \end{aligned} \tag{41}$$

All query branches, including a wrong proxy, enter this expression. The proxy is useful only if the resulting net column improves the relevant budget price. A pass on the prefix is also partial evidence, but its failure is an exact result. Neither a predicted pass nor a passed prefix alone certifies complete-suite success.

The learning assumptions hold under the following conditions. If all items are independent draws from this same six-atom law, every component is a core copy. Then price-region equality holds by replication for every law, and type labels are the common protocol label. If every atom is at least $p_{\min} > 0$, every unresolved terminal posterior has both binary probabilities at least $p_{\min}$. A deterministic partial policy with positive gain retains both outcomes on at least one terminal history; let ω_1, ω_0 be compatible atoms with respective outcomes $1, 0$. Since the contributions are binary,

$$\begin{aligned} \Pr(X = z \mid \mathfrak{h}) \geq \frac{\mu(\omega_z)}{\Pr(\mathfrak{h})} \geq p_{\min} \quad (z = 0, 1), \\ B^+_\pi = \mathbb{E}[X - l(\mathfrak{h}_\pi)] \geq \mu(\omega_1) \geq p_{\min}, \\ B^-_\pi = \mathbb{E}[u(\mathfrak{h}_\pi) - X] \geq \mu(\omega_0) \geq p_{\min}. \end{aligned}$$

Thus all positive deterministic columns meet $b_0 = \beta_0 = p_{\min}$ and generate the region. The complete routine above is common to the entire family.

Independent complete calibration records give a finite-confidence certificate for the mass floor. For m records of the full triple, write $\widehat{p}_i$ for the six empirical frequencies and, for $0 < \zeta < 1$, set

$$t_m = \sqrt{\frac{\log(12/\zeta)}{2m}}, \qquad \underline{p}_m = \max\{0, \min_{i \leq 6} \widehat{p}_i - t_m\}. \tag{42}$$

For independent complete records,

$$\begin{aligned} \Pr\{\min_i p_i < \underline{p}_m\} &\leq \sum_{i=1}^{6} \Pr\{\widehat{p}_i - p_i > t_m\} \\ &\overset{\text{Hoeffding}}{\leq} 6e^{-2mt_m^2} = 6e^{-\log(12/\zeta)} = \zeta/2 \leq \zeta. \end{aligned}$$

On a positive certificate the displayed coverage bounds hold for a fresh batch of the same law. Acquiring calibration triples costs their full query fees, including proxies. The fourth-root theorem takes the resulting coverage constants as input. Its cost guarantee requires independent draws from the same law. The output certificate always uses the full logical support.

An implementation of a fixed evaluator must return the prescribed actions and release category. The six-atom example gives one concrete interface for this requirement. The experiments instantiate the query laws and fees stated in Appendix E.2.

B.5 A CATALOG GENERATING THE LOCAL PRICES

For a fixed core let $\Pi_{0,+}$ be a declared positive-gain policy library and $\mathcal{D}_0 = \{\lambda \geq 0 : \lambda \cdot B_\pi \geq G_\pi \ (\pi \in \Pi_{0,+})\}$ its dual region. Choose a finite subcatalog $\mathcal{C} \subseteq \Pi_{0,+}$ whose inequalities define that same region:

$$\mathcal{D}_{\mathcal{C}} = \{\lambda \geq 0 : \lambda \cdot B_\pi \geq G_\pi \ (\pi \in \mathcal{C})\} = \mathcal{D}_0, \qquad \beta_{\text{gen}} = \min_{\pi \in \mathcal{C},\ s \in \{+,-\}} B^s_\pi. \tag{43}$$

Finite linear-programming duality gives $\mathcal{V}_{\mathcal{C}}(c) = \mathcal{V}_0(c)$ for every $c \geq 0$, also with an outer side omitted. Redundant inequalities and duplicate rays may therefore be removed before defining β_{gen}.

Fix this subcatalog once, and choose all subsequent basic solutions from it. The requirement is the existence of one such generating subcatalog, not that every positive policy or every alternative optimum have large exposure. The library may be all positive policies of the core, but need not be: unrepresented policies remain in the optimal comparator and are accounted for by (65). "Fixed" in a conditional execution means chosen before its responses; laws and library data may vary between batch sizes when the displayed geometric and execution bounds remain uniform.

For any chosen optimal allocation z and any present budget c_s,

$$\beta_{\mathrm{gen}} \sum_{\pi \in \mathcal{C}} z_\pi \le \sum_{\pi \in \mathcal{C}} B^s_\pi z_\pi \le c_s, \qquad \sum_\pi z_\pi \le d_\Gamma(x)/\beta_{\mathrm{gen}}. \tag{44}$$

A basic optimum has at most two nonzero columns. Thus the reserve feasibility, activation probabilities and exit allocation use only this exposure bound. The optimal comparator retains the complete legal query class.

For a chosen two-column basis $\mathbf{Q} = [B_1, B_2]$, differentiate $z(x) = \mathbf{Q}^{-1}(x-a, b-x)^T$ to obtain

$$z_1'(x) = \frac{B_2^+ + B_2^-}{\det \mathbf{Q}}, \qquad z_2'(x) = -\frac{B_1^+ + B_1^-}{\det \mathbf{Q}}.$$

For a one-column basis on side s, the coefficient slope has magnitude $1/B^s_\pi$; an existing other slack has magnitude $1 + B^{-s}_\pi / B^s_\pi$. Only finitely many nonsingular chosen bases occur. Define

$$\kappa_{\mathrm{bas}} = \max_{\mathbf{Q}\ \text{used}} \|\mathbf{Q}^{-1}\|_\infty, \qquad a_{\mathrm{bas}} = \min_{\text{nonzero coefficient/slack slopes used}} |\text{slope}| > 0. \tag{45}$$

One-side reciprocals are included in the corresponding raw increment bounds. These finite quantities, the core completion prices and β_{gen} determine the tracking constants. The constants grow with the inverse norm of a nearly singular basis.

The choice of catalog matters even when the price is unchanged. Append an independent nuisance bit $Z \sim \mathrm{Bernoulli}(\varepsilon)$ of fee ε^2 to the $(r, c_h, \varepsilon_{\mathrm{h}}) = (10, 1, 1/10)$ example. The pure policy that queries Z, retains Y only on $Z = 1$, and otherwise reads Y has

$$(B^+, B^-, G) = (\varepsilon/2, \varepsilon/2, 10\varepsilon - \varepsilon^2).$$

For $0 < \varepsilon < 1$ it has positive gain, so the all-policy exposure minimum is at most $\varepsilon/2$. Yet every use of Z can be simulated by a private coin of the same law, with the same subsequent observations and lower fee. The original three columns still generate the entire value. For them $\beta_{\mathrm{gen}} = 1/20$, independently of ε. The selected query kernels and their reachable recourse histories are unchanged, so the cost controller's constants are unchanged as well. This is a statement about chosen cost programs; a learner exploring the enlarged graph has its own explicit graph-complexity term.

B.6 Boundary exposure and cost balance

For $I_k = \mathbf{1}\{K \ge k\}$ put

$$D_k^+ = \mathbb{E}(K-k)_+, \quad D_k^- = \mathbb{E}(k-1-K)_+, \quad D_k = D_k^+ + D_k^-. \tag{46}$$

At joint termination let $l = \min K(\Omega_{\mathfrak{h}})$, $u = \max K(\Omega_{\mathfrak{h}})$, $m = \mathbb{E}[K \mid \mathfrak{h}]$, $o_+ = I_k(l-k)$, $o_- = (1-I_k)(k-1-u)$. Exactness makes I_k measurable in $\mathfrak{h}$. Thus

$$\begin{aligned} \mathbb{E}[(K-k)_+ \mid \mathfrak{h}] &= \mathbb{E}[I_k(K-k) \mid \mathfrak{h}] = I_k(m-k) = I_k(m-l) + o_+, \\ \mathbb{E}[(k-1-K)_+ \mid \mathfrak{h}] &= \mathbb{E}[(1-I_k)(k-1-K) \mid \mathfrak{h}] \\ &= (1-I_k)(k-1-m) = (1-I_k)(u-m) + o_-. \end{aligned} \tag{47}$$

Taking expectations gives

$$\mathbb{E}[I_k(m-l)] + \mathbb{E}o_+ = D_k^+, \qquad \mathbb{E}[(1-I_k)(u-m)] + \mathbb{E}o_- = D_k^-. \tag{48}$$

Under (P), set $V_t = \sum_g v_g(\mathfrak{h}_g^t)$ and $\mathscr{E}_\pi = \sum_{t<T} d_{g_t}(\mathfrak{h}_{g_t}^t, e_t)$. Only component g_t changes at query t, so

$$\begin{aligned} \mathbb{E}[V_{t+1} - V_t \mid \mathscr{F}_t] &= \mathbb{E}[v_{g_t}(\mathfrak{h}_{g_t}^{t+1}) \mid \mathscr{F}_t] - v_{g_t}(\mathfrak{h}_{g_t}^t) \\ &\overset{(27)}{=} d_{g_t}(\mathfrak{h}_{g_t}^t, e_t) - w_{e_t}, \\ \mathbb{E}C_\pi &= \mathbb{E}\sum_{t<T} d_{g_t}(\mathfrak{h}_{g_t}^t, e_t) - \mathbb{E}\sum_{t<T}(V_{t+1} - V_t) \\ &= V_0 - \mathbb{E}V_T + \mathbb{E}\mathscr{E}_\pi. \end{aligned} \tag{49}$$

At most $\sum_g |E_g|$ distinct responses are queried. Product compatibility at termination also gives

$$l = \sum_g l_g, \quad u = \sum_g u_g, \quad m = \sum_g m_g, \quad m - l = \sum_g b_g^+, \quad u - m = \sum_g b_g^-. \tag{50}$$

Since $v_g \le \mathfrak{c}_s b_g^s$ on either certified side,

$$\begin{aligned}\mathbb{E}V_T &\le \mathfrak{c}_+ \mathbb{E}[I_k(m-l)] + \mathfrak{c}_- \mathbb{E}[(1-I_k)(u-m)] \\ &\overset{(48)}{=} \mathfrak{c}_+(D_k^+ - \mathbb{E}o_+) + \mathfrak{c}_-(D_k^- - \mathbb{E}o_-).\end{aligned}$$

Thus (49) becomes

$$\mathbb{E}C_\pi \ge V_0 - \mathfrak{c}_+ D_k^+ - \mathfrak{c}_- D_k^- + \mathfrak{c}_+ \mathbb{E}o_+ + \mathfrak{c}_- \mathbb{E}o_- + \mathbb{E}\mathscr{E}_\pi. \tag{51}$$

The net bound keeps the relationship between the last term and retained work.

B.7 Component-excluding prices and the lower bound

Lemma A (Component-excluding prices)**.** *Assume (P) and bounded integer aggregates $0 \le X_g \le x_{\max}$. An empty positive-gain menu gives $C_N^* = V_0$. Otherwise write $\Lambda = \widehat{\rho}_+ + \widehat{\rho}_-$ and $b_* = \min_{g,\mathfrak{h}\in\mathcal{S}_g,s} b_g^s(\mathfrak{h}) > 0$. There is a piecewise-constant optimal dual selection $\lambda_N(x)$ with jump set Ξ_N such that, for $q_\Gamma \ge 1$ genuine count boundaries,*

$$\alpha_\xi = \|\lambda_N(\xi+) - \lambda_N(\xi-)\|_1, \qquad \sum_{\xi\in\Xi_N} \alpha_\xi \le (2q_\Gamma - 1)\Lambda. \tag{52}$$

For a genuine boundary k put

$$\mathcal{A}_\xi = \{|K-\xi| \le x_{\max}\}, \quad \chi_\xi = \Pr(\mathcal{A}_\xi), \quad D_k = \mathbb{E}[(K-k)_+ + (k-1-K)_+]. \tag{53}$$

Every exact adaptive query policy satisfies

$$\mathbb{E}C_\pi \ge V_0 - \mathbb{E}\mathcal{F}^{\mathrm{net}}(K) - \frac{2x_{\max}}{b_*} \sum_\xi \alpha_\xi \left[(|\xi - k| + x_{\max})\chi_\xi + \frac{x_{\max}}{2}\sqrt{D_k \chi_\xi / b_*} \right]. \tag{54}$$

Under (K), $b_* \ge b_0 > 0$ and $\Lambda \le \Lambda_0$, the correction is $O(1)$ with constants depending only on $x_{\max}, b_0, \Lambda_0, \nu, c_{\mathrm{gate}}, q_0$. The number of local price vertices may vary with N. For binary aggregates, let w_Γ be the largest width $b-a$ of an internal count cell (zero if there is none), and let Ξ_N^{int} contain only price changes strictly inside such cells. Then, without an exposure floor or (K),

$$C_N^* \ge V_0 - \mathbb{E}\mathcal{F}^{\mathrm{net}}(K) - 2w_\Gamma \sum_{\xi\in\Xi_N^{\mathrm{int}}} \alpha_\xi \Pr(|K-\xi| \le 1). \tag{55}$$

More generally, if every component satisfies (SA),

$$C_N^* \ge V_0 - \mathbb{E}\mathcal{F}^{\mathrm{net}}(K) - 2x_{\max} w_\Gamma \sum_{\xi\in\Xi_N^{\mathrm{int}}} \alpha_\xi \Pr(|K-\xi| \le x_{\max}). \tag{56}$$

Under (K) and bounded prices both corrections are $O(1)$, regardless of how small the background posteriors become. For one threshold the sum is empty, giving $C_N^* \ge V_0 - S_N$ for all support-anchored aggregates at every finite N.

Proof. Component-excluding projection. Condition on all full mechanisms ω_{-g} except component g and simulate their responses inside the global policy. By (P), this induces a randomized legal local policy under its original law μ_g. Truncating queries after identification of (A_g, X_g) decreases its fee and leaves zero exposure. For fixed ω_{-g}, let $q_{\pi'|\omega_{-g}}$ be the induced local mixture. Positive-gain branches satisfy the dual constraint. Other branches satisfy $G_{\pi'} \le 0 \le \lambda \cdot B_{\pi'}$. For every feasible ω_{-g}-measurable price,

$$\begin{aligned}v_g(\mathfrak{h}_g^0) - \mathbb{E}[C_{\pi,g} \mid \omega_{-g}] &\le \sum_{\pi'} q_{\pi'|\omega_{-g}} G_{\pi'} \\ &\le \sum_{\pi'} q_{\pi'|\omega_{-g}} \lambda_g(\omega_{-g}) \cdot B_{\pi'} \\ &= \lambda_g(\omega_{-g}) \cdot \mathbb{E}[b_g(\mathfrak{h}_g^T) \mid \omega_{-g}].\end{aligned} \tag{57}$$

If the positive menu is empty, use $\lambda_g = 0$ and sum to obtain $C_N^* = V_0$.

Price variation on count cells. On an internal cell (a, b) select Pareto-efficient optimal dual vertices. For $x < x'$, with $c(x) = (x - a, b - x)$, set $d = \lambda_N(x') - \lambda_N(x)$. Optimality gives

$$c(x) \cdot d \geq 0, \qquad c(x') \cdot d \leq 0, \qquad d_+ - d_- \leq 0.$$

The signs follow by contradiction:

$$\begin{aligned} d_+ > 0 &\implies d_- \geq d_+ > 0 \implies c(x') \cdot d > 0, \\ d_- < 0 &\implies d_+ \leq d_- < 0 \implies c(x) \cdot d < 0. \end{aligned}$$

Thus $d_+ \leq 0 \leq d_-$. The variation within a positive-width cell is at most Λ, including axial choices $(\widehat{\rho}_+, 0)$ at its left end and $(0, \widehat{\rho}_-)$ at its right end. A singleton cell uses either axis; each outer cell has a constant axial price. Extend the rightmost price of a cell constantly across the next unit gap. There are $q_\Gamma - 1$ internal cells and q_Γ intervening transitions, each of variation at most Λ, proving (52). For $\vartheta_N = \lambda_{N,+} - \lambda_{N,-}$,

$$\|\lambda_N(x) - \lambda_N(y)\|_\infty, \quad |\vartheta_N(x) - \vartheta_N(y)| \leq \sum_{\xi \in [x \wedge y, x \vee y]} \alpha_\xi. \tag{58}$$

Prices extend constantly beyond the outer endpoints.

Terminal-posterior centering. At termination let $m_g = \mathbb{E}[X_g \mid \mathscr{F}_T]$, $m = \sum_g m_g$, and let R_T count unresolved component aggregates. Apply (57) with $\lambda_g = \lambda_N(K - X_g + \mathbb{E}X_g)$. For an unresolved component, $|X_g - \mathbb{E}X_g| \leq x_{\max}$ and (58) give

$$(\lambda_g - \lambda_N(K)) \cdot b_g \leq x_{\max} \sum_\xi \alpha_\xi \mathbf{1}_{\mathcal{A}_\xi}.$$

For a resolved component $b_g = 0$. On a record certifying $[a, b]$, put $L_T = \sum_g l_g$ and $U_T = \sum_g u_g$. Since $L_T \geq a$ and $U_T \leq b$, the nonnegative prices give

$$\begin{aligned} \lambda_N(K) \cdot \sum_g b_g &= \lambda_{N,+}(K)(m - L_T) + \lambda_{N,-}(K)(U_T - m) \\ &\leq \lambda_{N,+}(K)(m - a) + \lambda_{N,-}(K)(b - m) \\ &= \mathcal{F}^{\mathrm{net}}(K) + \vartheta_N(K)(m - K). \end{aligned}$$

An absent outer side has zero price. Summing (57), then adding and subtracting the common price, gives

$$\begin{aligned} V_0 - \mathbb{E}C_\pi &\leq \sum_g \mathbb{E}[\lambda_g \cdot b_g] \\ &= \mathbb{E}\left[\lambda_N(K) \cdot \sum_g b_g\right] + \mathbb{E}\sum_g (\lambda_g - \lambda_N(K)) \cdot b_g. \end{aligned}$$

The endpoint bound and (58) now give

$$V_0 - \mathbb{E}C_\pi \leq \mathbb{E}\mathcal{F}^{\mathrm{net}}(K) + \mathbb{E}[\vartheta_N(K)(m - K)] + x_{\max} \sum_\xi \alpha_\xi \mathbb{E}[R_T \mathbf{1}_{\mathcal{A}_\xi}]. \tag{59}$$

Let $\mathscr{F}_T$ and $\mathscr{G}_{g,T}$ be the sigma-fields in Appendix A.1. The first projection price $\lambda_N(K - X_g + \mathbb{E}X_g)$ is $\sigma(\omega_{-g})$-measurable. After the projection has been summed, the centering multiplier $\vartheta_N(K - X_g + m_g)$ is instead $\mathscr{G}_{g,T}$-measurable: $K - X_g$ is known from ω_{-g} and m_g from $\mathscr{F}_T$. Equation (23) therefore gives

$$\mathbb{E}[(m_g - X_g)\vartheta_N(K - X_g + m_g) \mid \mathscr{G}_{g,T}] = 0. \tag{60}$$

Subtract this identity componentwise. Since $|X_g - m_g| \leq x_{\max}$,

$$\begin{aligned} \mathbb{E}[\vartheta_N(K)(m - K)] &= \mathbb{E}\sum_g (m_g - X_g)[\vartheta_N(K) - \vartheta_N(K - X_g + m_g)] \\ &\leq x_{\max} \sum_\xi \alpha_\xi \mathbb{E}[R_T \mathbf{1}_{\mathcal{A}_\xi}]. \end{aligned} \tag{61}$$

For any genuine k, put $D(K) = (K-k)_+ + (k-1-K)_+$. On a certificate above k, $b_* R_T \le m - L_T \le m - k$. On a certificate below k, $b_* R_T \le U_T - m \le k - 1 - m$. Consequently

$$\begin{aligned} b_* R_T &\le \mathbb{E}[D(K) \mid \mathscr{F}_T] \le |m-k|, \\ b_* \mathbb{E} R_T &\le \mathbb{E} D(K) = D_k, \\ \mathbb{E}(K-m)^2 = \mathbb{E}\operatorname{Var}(K \mid \mathscr{F}_T) &= \mathbb{E}\sum_g \operatorname{Var}(X_g \mid \mathscr{F}_T) \le \tfrac{x_{\max}^2}{4}\mathbb{E} R_T. \end{aligned}$$

The first inequality and Cauchy–Schwarz now give

$$\begin{aligned} b_* \mathbb{E}[R_T \mathbf{1}_{\mathcal{A}_\xi}] &\le \mathbb{E}[|m-k| \mathbf{1}_{\mathcal{A}_\xi}] \\ &\le \mathbb{E}[(|K-\xi| + |\xi - k| + |K-m|)\mathbf{1}_{\mathcal{A}_\xi}] \\ &\le (x_{\max} + |\xi - k|)\chi_\xi + \sqrt{\mathbb{E}(K-m)^2 \chi_\xi} \\ &\le (x_{\max} + |\xi - k|)\chi_\xi + \frac{x_{\max}}{2}\sqrt{D_k \chi_\xi / b_*}. \end{aligned} \tag{62}$$

$$\begin{aligned} V_0 - \mathbb{E}C_\pi - \mathbb{E}\mathcal{F}^{\mathrm{net}}(K) &\overset{(61)}{\le} 2x_{\max} \sum_\xi \alpha_\xi \mathbb{E}[R_T \mathbf{1}_{\mathcal{A}_\xi}] \\ &\overset{(62)}{\le} \frac{2x_{\max}}{b_*} \sum_\xi \alpha_\xi \left[(|\xi - k| + x_{\max})\chi_\xi + \frac{x_{\max}}{2}\sqrt{D_k \chi_\xi / b_*} \right]. \end{aligned}$$

Heterogeneous lattice concentration. Let $p_g(x) = \Pr(X_g = x)$ and $Q_{x_{\max}} = \sum_{0 \le a < b \le x_{\max}} (b-a)^2$. Then

$$\sigma_N^2 = \sum_{a<b} (b-a)^2 \sum_g p_g(a) p_g(b) \le Q_{x_{\max}} \max_{a<b} \sum_g \min\{p_g(a), p_g(b)\}.$$

Choose a maximizing pair. Independently split each law into probability $r_g = 2\min\{p_g(a), p_g(b)\}$ on a fair $\{a,b\}$ draw and its residual law. For the independent gate count M, $\bar{M} = \mathbb{E}M \ge 2\sigma_N^2 / Q_{x_{\max}}$ and

$$K \mid (\text{gates and residual draws}) \overset{d}{=} k_0 + (b-a)\operatorname{Binomial}(M, 1/2).$$

Write $\bar{M} = \mathbb{E}M$. The Bernoulli Chernoff bound and the binomial point-mass bound are $\Pr(M \le \bar{M}/2) \le e^{-\bar{M}/8}$ and $\max_j 2^{-m}\binom{m}{j} \le \sqrt{2/(m+1)}$. Thus

$$\begin{aligned} \sup_z \Pr(K = z) &\le \Pr(M \le \bar{M}/2) + \mathbb{E}[\sqrt{2/(M+1)}\, \mathbf{1}_{\{M > \bar{M}/2\}}] \\ &\le e^{-\bar{M}/8} + \sqrt{2/(\bar{M}/2 + 1)} \\ &\le e^{-\sigma_N^2/(4Q_{x_{\max}})} + \sqrt{2Q_{x_{\max}}/(\sigma_N^2 + Q_{x_{\max}})}. \end{aligned}$$

Equivalently,

$$\sup_z \Pr(K = z) \le e^{-\sigma_N^2/(4Q_{x_{\max}})} + \sqrt{2Q_{x_{\max}}/(\sigma_N^2 + Q_{x_{\max}})}. \tag{63}$$

General bounds on concentration functions are given by Juškevičius (2024). The two-point decomposition above applies to heterogeneous component laws. Under (K), $\chi_\xi = O(N^{-1/2})$ uniformly, all price changes lie within $O(\sqrt{N})$ of k, and $D_k = O(\sqrt{N})$. The variation budget (52) makes the correction bounded independently of the number of price vertices.

Compatible segments replace a posterior floor. Assume (SA), choose a fixed a_g^{anc} from its intersection, and in (57) use $\lambda_g = \lambda_N(K - X_g + a_g^{\mathrm{anc}})$. This price is measurable in the other complete mechanisms, as required by that projection. Fix an unresolved component g at a terminal transcript certifying $[a,b]$, and put $S_g = K - X_g$. Product compatibility and the interval-valued output imply

$$[S_g + l_g(\mathfrak{h}_g^T), S_g + u_g(\mathfrak{h}_g^T)] \subseteq [a,b], \qquad S_g + a_g^{\mathrm{anc}},\ S_g + m_g,\ K \in [a,b].$$

The anchor is a fixed point in the current support hull; it need not equal $\mathbb{E}X_g$ or be itself an attained outcome. Consequently its exclusion-price argument differs from K by at most $x_{\max}$. The terminal

centering identity (60) still uses m_g and is unchanged. Thus both price differences in (59) and (61) cross only internal price knots, never a genuine output boundary. On an outer cell the price is constant; on a resolved component both residual exposure and centered fluctuation are zero. On an internal cell, integer-valued product support gives

$$R_T \le \sum_g (u_g - l_g) = u - l \le b - a \le w_\Gamma.$$

For each unresolved component, $b_g^+ + b_g^- \le x_{\max}$ and $|m_g - X_g| \le x_{\max}$. The same projection and centering identities therefore give

$$V_0 - \mathbb{E}C_\pi - \mathbb{E}\mathcal{F}^{\mathrm{net}}(K) \le 2x_{\max} \sum_{\xi \in \Xi_N^{\mathrm{int}}} \alpha_\xi \mathbb{E}[R_T \mathbf{1}\{|K-\xi| \le x_{\max},\ \text{certified cell contains } \xi\}].$$

Since $R_T \le w_\Gamma$ on each indicated event,

$$V_0 - \mathbb{E}C_\pi - \mathbb{E}\mathcal{F}^{\mathrm{net}}(K) \le 2x_{\max} w_\Gamma \sum_{\xi \in \Xi_N^{\mathrm{int}}} \alpha_\xi \Pr(|K-\xi| \le x_{\max}).$$

This is (56). For binary aggregates, every unresolved support is $\{0,1\}$; take $a_g^{\mathrm{anc}} = 0$ and $x_{\max} = 1$ to obtain (55). Within each internal cell the variation is at most Λ, so $\sum_{\Xi_N^{\mathrm{int}}} \alpha_\xi \le (q_\Gamma - 1)\Lambda$. A knot window contains at most $2x_{\max}+1$ lattice points. Consequently the correction is bounded by

$$2x_{\max}(2x_{\max}+1) w_\Gamma (q_\Gamma - 1)\Lambda \sup_z \Pr(K = z) = O(1),$$

using $w_\Gamma \le 2c_{\mathrm{gate}}\sqrt{N}$ under (K) and (63).

A strict enlargement of endpoint preservation. Let an aggregate have outcomes $0, 1, 2$ and a marker with four positive joint masses

$$\Pr(X, Z) = (18/25, 2/25, 1/10, 1/10) \quad \text{on } (0,0), (1,0), (1,1), (2,1).$$

A marker query leaves support $\{0,1\}$ or $\{1,2\}$; a direct aggregate query resolves it. All actions are known. The anchor $a^{\mathrm{anc}} = 1$ belongs to every unresolved support hull. Both initial endpoints are removed on some history, and $\mathbb{E}X = 19/50$ is excluded on $Z = 1$. Thus this example satisfies (SA), but neither (EP) nor prior-mean preservation. Strictly positive noisy refinements before the direct query preserve these hulls and may have arbitrarily small posterior masses and growing depth.

Binary aggregates and one threshold. If $X_g \in \{0,1\}$ choose in (57)

$$\lambda_g = \begin{cases} (\widehat{\rho}_+, 0), & K - X_g \ge k, \\ (0, \widehat{\rho}_-), & K - X_g \le k-2, \\ \text{any feasible price}, & K - X_g = k-1. \end{cases}$$

In the last case, an unresolved X_g would admit both threshold outputs with all other actual components fixed; hence $b_g(\mathfrak{h}_g^T) = 0$. Every other unresolved component receives the certified side's price. Let $S_+ = \{K \ge k\}$ and $S_- = \{K \le k-1\}$. Then

$$\begin{aligned} V_0 - \mathbb{E}C_\pi &\le \mathbb{E}\sum_g \lambda_g \cdot b_g \\ &= \widehat{\rho}_+ \mathbb{E}[\mathbf{1}_{S_+} \sum_g b_g^+] + \widehat{\rho}_- \mathbb{E}[\mathbf{1}_{S_-} \sum_g b_g^-] \\ &\overset{(48)}{=} \widehat{\rho}_+ (D_k^+ - \mathbb{E}o_+) + \widehat{\rho}_- (D_k^- - \mathbb{E}o_-). \end{aligned}$$

Rearranging gives

$$\mathbb{E}C_\pi \ge V_0 - \widehat{\rho}_+ D_k^+ - \widehat{\rho}_- D_k^- + \widehat{\rho}_+ \mathbb{E}o_+ + \widehat{\rho}_- \mathbb{E}o_-. \tag{64}$$

This is finite in N and uses no criticality or common exposure floor. □

A library comparison with heterogeneous shortfalls. The same argument can price a chosen local library without requiring its dual region to equal the batch region. Let $\mathcal{D}_0$ be that library's region, $\mathcal{P}_0$ its vertices, F_0 its count envelope, and Λ_0 the sum of its axial prices. For each component define

$$d_g = \max_{\lambda \in \mathcal{P}_0} [v_g(\mathfrak{h}_g^0) - U_{g,\lambda}(\mathfrak{h}_g^0)], \qquad \mathfrak{D}_N = \sum_{g \in \mathcal{A}} d_g + \sum_{g \notin \mathcal{A}} v_g(\mathfrak{h}_g^0), \tag{65}$$

where $\mathcal{A}$ is any declared set of components satisfying (SA). The priced value U is defined in (35). Full evaluation is feasible and all terminal prices are nonnegative, so $0 \le d_g \le v_g(\mathfrak{h}_g^0)$. Let $\mathcal{I}_\Gamma$ contain the internal count cells, let $\Xi_0(I)$ be the library price knots strictly inside $I = [a_I, b_I]$. Put $\mathfrak{p}_N = \sup_z \Pr(K = z)$ and $W_N = \sum_{I \in \mathcal{I}_\Gamma} (b_I - a_I)$. Then

$$\begin{aligned}
\mathfrak{B}_N &= 2x_{\max} \sum_{I \in \mathcal{I}_\Gamma} (b_I - a_I) \sum_{\xi \in \Xi_0(I)} \alpha_\xi \Pr(K \in I, |K - \xi| \le x_{\max}), \\
C_N^* &\ge V_0 - \mathbb{E}F_0(K) - \mathfrak{D}_N - \mathfrak{B}_N, \\
\mathfrak{B}_N &\le 2x_{\max} \Lambda_0 \sum_{I \in \mathcal{I}_\Gamma} (b_I - a_I) \Pr(K \in I), \\
\mathfrak{B}_N &\le 2x_{\max}(2x_{\max} + 1) \Lambda_0 \mathfrak{p}_N W_N.
\end{aligned} \tag{66}$$

An empty internal-cell sum is zero. There is no restriction on the number or location of the thresholds, and no global posterior floor.

To prove this, first give the complete mechanisms outside $\mathcal{A}$ to a comparison oracle at zero fee. Simulating those queries for free can only decrease optimal cost. Conditional on their mechanisms, the remaining components retain their product law; their full baseline is $V_0 - \sum_{g \notin \mathcal{A}} v_g(\mathfrak{h}_g^0)$. For $g \in \mathcal{A}$, local priced optimality gives

$$v_g(\mathfrak{h}_g^0) - \mathbb{E}[C_{\pi,g} \mid \omega_{-g}] \le \lambda_g \cdot \mathbb{E}[b_g(\mathfrak{h}_g^T) \mid \omega_{-g}] + d_g$$

whenever λ_g is a library vertex selected using ω_{-g}. Use $\lambda_g = \lambda_0(K - X_g + a_g^{\mathrm{anc}})$ in the projection and $\vartheta_0(K - X_g + m_g)$ in terminal centering, exactly as in (60). The revealed components have no residual exposure or centered fluctuation. Both price arguments for any remaining unresolved component stay in its certified cell. On a transcript certifying I, integer product support gives

$$R_T \le \sum_{g \in \mathcal{A}} (u_g - l_g) \le b_I - a_I.$$

Let C_{free}^* denote the free-revelation oracle cost and $V_{\mathcal{A}} = \sum_{g \in \mathcal{A}} v_g(\mathfrak{h}_g^0)$. The full projection and centering sum is

$$\begin{aligned}
&V_{\mathcal{A}} - C_{\mathrm{free}}^* - \mathbb{E}F_0(K) - \sum_{g \in \mathcal{A}} d_g \\
&\quad \le 2x_{\max} \sum_{I \text{ internal}} \sum_{\xi \in \Xi_0(I)} \alpha_\xi \mathbb{E}[R_T \mathbf{1}_{\{K \in I,\ |K - \xi| \le x_{\max}\}}] \\
&\quad \le 2x_{\max} \sum_I (b_I - a_I) \sum_{\xi \in \Xi_0(I)} \alpha_\xi \Pr(K \in I,\ |K - \xi| \le x_{\max}) = \mathfrak{B}_N, \\
&C_N^* \ge C_{\mathrm{free}}^* \ge V_0 - \mathbb{E}F_0(K) - \mathfrak{D}_N - \mathfrak{B}_N.
\end{aligned}$$

Coordinate monotonicity gives $\sum_{\xi \in \Xi_0(I)} \alpha_\xi \le \Lambda_0$. Each window contains at most $2x_{\max} + 1$ integers. Therefore

$$\begin{aligned}
\Pr(K \in I,\ |K - \xi| \le x_{\max}) &\le \min\{\Pr(K \in I), (2x_{\max} + 1)\mathfrak{p}_N\}, \\
\mathfrak{B}_N &\le 2x_{\max} \Lambda_0 \sum_I (b_I - a_I) \Pr(K \in I) \le 2x_{\max} \Lambda_0 w_\Gamma, \\
\mathfrak{B}_N &\le 2x_{\max}(2x_{\max} + 1) \Lambda_0 \mathfrak{p}_N \sum_I (b_I - a_I).
\end{aligned}$$

For anchored components, $d_g = 0$ exactly when every library vertex passes that component's priced Bellman test. A conservative certificate is $d_g \le v_g(\mathfrak{h}_g^0)$, or any known upper bound on a legal full routine's expected fee. Thus a small total fee of exceptional interfaces does not require their graphs or normalized price ratios to be controlled.

C EXECUTING THE ALLOCATION AND PROVING THE COST COMPARISON

The lower bound prices every legal policy. To obtain an upper bound, the allocation must fit the available components and end with an exact certificate. We first bound the cost of completing a partial certificate. We then select a reserve and adjust the allocation as responses arrive. The final two subsections bound insufficient supply and combine these estimates to prove Theorem 1.

C.1 RESERVE SELECTION AND EXACT COMPLETION

In the fixed core let $C_{\mathrm{loc}} = \max_j \sum_{e\in E_j} w_e$, $\bar v = \max_{j,\mathfrak{h}} v_j(\mathfrak{h})$ and $b_* = \min_{j,\mathfrak{h}\in\mathcal{S}_j,s} b_j^s(\mathfrak{h}) > 0$. A whole local completion costs at most C_{loc}. The sum of address fees is one sufficient choice; any uniform bound on the chosen partial/full executions and their recourse may replace it. Unqueried nuisance addresses need not be included in that bound.

Lemma B (Exact endpoint recourse)**.** *At a history where all actions are known, let $[L, U]$ be the attainable count interval and $[a, b]$ any classifier cell. Use minimizing full continuations, and let $\mathfrak{c}_s = \max_{g,\mathfrak{h}\in\mathcal{S}_g} v_g(\mathfrak{h})/b_g^s(\mathfrak{h})$ be the directional completion prices of* (28)*. Under (P), a legal policy identifies $\Gamma(K)$ at conditional expected extra cost at most*

$$\mathfrak{c}_+\big[(a-L)_+ + x_{\max}\mathbf{1}\{L<a\}\big] + \mathfrak{c}_-\big[(U-b)_+ + x_{\max}\mathbf{1}\{U>b\}\big]. \tag{67}$$

Absent outer deficits are zero. This holds for each finite heterogeneous system with positive support, using its own prices and $x_{\max} = \max_g x_{\max,g}$. In particular, $\mathfrak{c}_s \le \bar v/b_* \le C_{\mathrm{loc}}/b_*$ recovers

$$\frac{C_{\mathrm{loc}}}{b_*}\big[(a-L)_+ + (U-b)_+ + 2x_{\max}\big], \tag{68}$$

with C_{loc} replaceable by $\bar v = \max_{g,\mathfrak{h}} v_g(\mathfrak{h})$. The actual K need not lie in $[a, b]$. The bound (68) *also holds for any legal full continuations of pathwise fee at most C_{loc}, even if their expectations are not minimizing. Such continuations can be selected from the support graph without knowing the probabilities. More generally, choose support-valid continuations $\kappa_g(\mathfrak{h})$ from the graph and fees, and suppose*

$$\mathbb{E}_\mu[C_{\kappa_g(\mathfrak{h})} \mid \mathfrak{h}] \le \chi_{\mathrm{rec}} b_g^s(\mathfrak{h}), \qquad s\in\{+,-\}, \tag{69}$$

at every unresolved history where they are used. Their extra fee is at most $\chi_{\mathrm{rec}}[(a-L)_+ + (U-b)_+ + 2x_{\max}]$. This condition controls the expected price of completion rather than requiring a fixed lower bound on every posterior exposure.

Proof. While $L < a$ and a component is unresolved, complete it with its minimizing full continuation. Its conditional mean fee and endpoint increment satisfy

$$\mathbb{E}[C_g \mid \mathfrak{h}_t] = v_g(\mathfrak{h}_g) \le \mathfrak{c}_+ b_g^+(\mathfrak{h}_g),$$
$$0 \le L_{t+1} - L_t = X_g - l_g \le x_{\max}, \qquad \mathbb{E}[L_{t+1} - L_t \mid \mathfrak{h}_t] = b_g^+(\mathfrak{h}_g).$$

Let τ_+ be exhaustion or the first time $L_t \ge a$. With $D_+ = (a - L_{\mathrm{init}})_+$, bounded increments give

$$\begin{aligned}
L_{\tau_+} - L_{\mathrm{init}} &\le D_+ + x_{\max}\mathbf{1}_{\{L_{\mathrm{init}}<a\}},\\
\mathbb{E}[C_+ \mid \mathfrak{h}_{\mathrm{init}}] &= \sum_{t=0}^{N-1} \mathbb{E}[\mathbf{1}_{\{t<\tau_+\}}\mathbb{E}(C_{g_t} \mid \mathscr{F}_t) \mid \mathfrak{h}_{\mathrm{init}}]\\
&\le \mathfrak{c}_+ \sum_{t=0}^{N-1} \mathbb{E}[\mathbf{1}_{\{t<\tau_+\}}\mathbb{E}(L_{t+1} - L_t \mid \mathscr{F}_t) \mid \mathfrak{h}_{\mathrm{init}}]\\
&\overset{\text{tower}}{=} \mathfrak{c}_+\mathbb{E}[L_{\tau_+} - L_{\mathrm{init}} \mid \mathfrak{h}_{\mathrm{init}}]\\
&\le \mathfrak{c}_+[D_+ + x_{\max}\mathbf{1}_{\{L_{\mathrm{init}}<a\}}].
\end{aligned}$$

For the second phase use $\Delta U_t = U_t - U_{t+1} \ge 0$ and $\mathbb{E}[\Delta U_t \mid \mathscr{F}_t] = b_{g_t}^-$. Its initial upper endpoint satisfies $U_{\tau_+} \le U_{\mathrm{init}}$, so

$$\begin{aligned}
\mathbb{E}[C_- \mid \mathfrak{h}_{\tau_+}] &\le \mathfrak{c}_-[(U_{\tau_+} - b)_+ + x_{\max}\mathbf{1}_{\{U_{\tau_+}>b\}}]\\
&\le \mathfrak{c}_-[(U_{\mathrm{init}} - b)_+ + x_{\max}\mathbf{1}_{\{U_{\mathrm{init}}>b\}}],\\
\mathbb{E}[C_+ + C_- \mid \mathfrak{h}_{\mathrm{init}}] &\le \mathfrak{c}_+[D_+ + x_{\max}\mathbf{1}_{\{L_{\mathrm{init}}<a\}}]\\
&\quad + \mathfrak{c}_-[(U_{\mathrm{init}} - b)_+ + x_{\max}\mathbf{1}_{\{U_{\mathrm{init}}>b\}}].
\end{aligned}$$

Termination identifies all results or puts their interval inside $[a, b]$. The attainable support certifies the returned output. For a nonoptimal continuation with path fee at most C_{loc}, $\mathbb{E}[C_g \mid \mathscr{F}_t] \le C_{\mathrm{loc}} \le (C_{\mathrm{loc}}/b_*)b_g^s$. For the completion-price condition the first inequality is instead $\mathbb{E}[C_g \mid \mathscr{F}_t] \le \chi_{\mathrm{rec}} b_g^s$. In either case the displayed stopped sums give

$$\mathbb{E}[C_+ + C_- \mid \mathfrak{h}_{\mathrm{init}}] \le \chi[(a - L_{\mathrm{init}})_+ + (U_{\mathrm{init}} - b)_+ + 2x_{\max}], \quad \chi \in \{C_{\mathrm{loc}}/b_*, \chi_{\mathrm{rec}}\}.$$

□

A conditional reserve, with arbitrary initial supply. Fix a finite menu of J local types with positive laws and a nonempty generating catalog $\mathcal{C}$ from (43). At a stopping history, all components outside a pool of m_0 untouched roots per menu type have been fully evaluated. Conditional on that history, the pool has its original product law. No lower bound on m_0/N is imposed. Put $\beta_* = \beta_{\mathrm{gen}} = \min_{\pi\in\mathcal{C}} \min(B_\pi^+, B_\pi^-) > 0$, $\bar{G} = \max_{\pi\in\mathcal{C}} G_\pi$, $\mu_N = \mathbb{E}K$, and $\sigma_N^2 = \operatorname{Var} K$. On each real count cell define

$$d_\Gamma(x) = \min\{x - a \ (a > 0),\ b - x \ (b < n)\}, \qquad c_* = 8/\beta_*, \quad C_* = 1 + c_* J x_{\max}. \tag{70}$$

Set $d_\Gamma = 0$ on unit gaps. It is globally one-Lipschitz, and for any genuine threshold k, $d_\Gamma(x) \le |x - k| + 1$.

Starting from this pool, let M_t be the posterior mean of the final count after t balanced rounds of completing one reserved root per type; $m_t = m_0 - t$. If initially $m_0 \le c_* d_\Gamma(M_0) + C_*$, use full completion. Otherwise stop these rounds at the first time

$$m_T \le c_* d_\Gamma(M_T) + C_*. \tag{71}$$

At that time solve the generating-catalog net LP using the observed cell capacities at M_T. A basic solution uses at most two policies. The feedback construction updates their execution probabilities after every balanced round. When the reserve is bounded in size, or M_T is in a unit gap, complete the remaining roots. Exactness is always tested on the actual attainable endpoints, with Lemma B providing a legal correction.

$$\begin{aligned} &\text{balanced full completion} \longrightarrow \text{reserve stopping rule,} \\ &\text{feedback partial-policy allocation} \longrightarrow \text{deficit correction.} \end{aligned} \tag{72}$$

Every branch is exact; the stopping rule uses only already completed outcomes.

Supply holds at the reserve stopping time. One round changes M_t by at most $Jx_{\max}$. If $T > 0$, the preceding strict inequality and the Lipschitz property imply

$$\begin{aligned} m_T &= m_{T-1} - 1 > c_* d_\Gamma(M_{T-1}) + C_* - 1 \\ &\ge c_* d_\Gamma(M_T) + C_* - 1 - c_* J x_{\max} = c_* d_\Gamma(M_T). \end{aligned} \tag{73}$$

Equation (44) bounds the total assigned mass by $d_\Gamma(M_T)/\beta_*$. Thus even two policies of one type use at most m_T roots. The process must stop by $m_t = 0$, and the displayed inequality prevents a positive-budget allocation from lacking roots. Conditional on its complete stopping transcript, all untouched roots retain their independent original laws.

On the exceptional initial-full branch define $m_T = 0$ and $M_T = K$ after completion. Conditional Jensen then gives, for any genuine k,

$$\mathbb{E}m_T \le \bar{m}_N := c_*[\sigma_N + |\mu_N - k| + 1] + C_*. \tag{74}$$

Indeed,

$$\begin{aligned} \mathbb{E}|M_T - \mu_N| &\le \sqrt{\mathbb{E}\{\mathbb{E}[K - \mu_N \mid \mathcal{F}_T]\}^2} \overset{\text{conditional Jensen}}{\le} \sqrt{\mathbb{E}(K - \mu_N)^2} = \sigma_N, \\ \mathbb{E}m_T &\le c_* \mathbb{E} d_\Gamma(M_T) + C_* \\ &\le c_*[\mathbb{E}|M_T - \mu_N| + |\mu_N - k| + 1] + C_* \le \bar{m}_N. \end{aligned}$$

Under (K), this expectation is $O(\sqrt{N})$. Lemma D separately bounds the loss from the initial-full branch using its actual core supply.

Shared accounting for fresh-root assignments. Whenever each root's partial policy is assigned before observing that root, the law of iterated expectations gives

$$\mathbb{E}C_{\text{before correction}} = V_0 - \mathbb{E} \sum_{\text{assigned partial roots } g} G_{\pi_g}. \tag{75}$$

This includes every branch of each chosen policy and applies to predictable randomized assignments. A static allocation creates fluctuations of size $\sqrt{m_T}$; the following feedback construction tracks them before stopping.

Procedure 1: Track **on a supplied core.** Input: known local laws, full routines, selected library $\mathcal{C}$, and m_0 untouched copies per type. Set $a_R = 8/\beta_{\text{gen}}$, $b_R = 1 + a_R J x_{\max}$.

1. Complete the background. With m untouched copies per type and known outside count O, compute $M = O + m\sum_j \mathbb{E}X_j$.
2. If $m_0 \le a_R d_\Gamma(M) + b_R$ initially, finish fully. Otherwise complete one component per type in balanced rounds until $m \le a_R d_\Gamma(M) + b_R$.
3. In the cell containing M, choose an optimal basis of at most two columns and its dual price. Keep this basis and price fixed during tracking. A bounded reserve or a unit-gap mean is completed fully.
4. Recompute its coefficients $z_{r,i}$ from actual endpoints and r remaining copies. Activate coefficient i once $z_{r,i}/r \ge 1/32$. Choose column i with probability $z_{r,i}/r$ thereafter, and zero before; unassigned components use full routines. Stop at the coefficient, slack, or small-reserve exits specified in Appendix C.2.
5. Round positive exit coefficients down, execute feasible assignments, complete other fresh components, and pay for retained-result queries until the attainable interval is inside one cell or all results are known.

C.2 Conditional delayed-activation tracking

Lemma C (Tracking a supplied local reserve)**.** *Fix J finite local types with positive laws, a nonempty net menu, and its constants c_*, C_* from* (70)*. Let $\mathscr{F}$ be the sigma-field generated by a stopping transcript and $\mathcal{E} \in \mathscr{F}$ an entry event. On $\mathcal{E}$, all outside components are fully evaluated and m untouched roots per menu type retain their independent original laws. With $M = \mathbb{E}[K \mid \mathscr{F}]$, assume*

$$c_* d_\Gamma(M) < m \le c_* d_\Gamma(M) + C_*. \tag{76}$$

Let $\mathcal{K}_N$ consist of the genuine boundary endpoints $k_s - 1, k_s$ and the menu's finitely many basis-change locations, and let $d_{\text{knot}} = \operatorname{dist}(M, \mathcal{K}_N)$. Artificial domain endpoints $0, n$ are omitted unless they are genuine boundary endpoints; the outer value pieces extend affinely there. There are constants $A, c_{\text{tr}} > 0$ and an exact conditional policy whose remaining cost $C_{\text{remaining}}$ satisfies

$$\mathbb{E}[C_{\text{remaining}} \mid \mathscr{F}] \le m \sum_{j=1}^{J} v_j(\mathfrak{h}_j^0) - \mathbb{E}[\mathcal{F}^{\text{net}}(K) \mid \mathscr{F}] + A(1 + \sqrt{m}\, e^{-c_{\text{tr}} {d_{\text{knot}}}^2/m}). \tag{77}$$

Constants depend only on the menu, its basis geometry, and $x_{\max}$. For any finite threshold family define $\mathfrak{p}_N = \sup_z \Pr(K = z)$ and $\mathfrak{K}_N = \sum_{\xi \in \mathcal{K}_N}(1 + d_\Gamma(\xi))$. Then

$$\mathbb{E}[\mathbf{1}_{\mathcal{E}} \sqrt{m}\, e^{-c_{\text{tr}} {d_{\text{knot}}}^2/m}] \le A'(1 + \mathfrak{p}_N \mathfrak{K}_N). \tag{78}$$

In particular, finitely many central thresholds and $\mathfrak{p}_N = O(N^{-1/2})$ imply a bounded correction. No assumption on the number, laws, or type frequencies of the completed outside components enters the conditional assertion. If w_Γ is the maximum internal-cell width (zero for a single threshold), the same expectation is also at most $A''(1 + \sqrt{w_\Gamma + 1})$, without any count-mass hypothesis.

Proof. Allocation and fixed-price accounting. Abbreviate $v_j = v_j(\mathfrak{h}_j^0)$. All conditional statements in this proof concern $\mathcal{E}$. Bounded reserves are completed. Otherwise choose the cell $[a, b]$ containing M, an optimal basic allocation from $\mathcal{C}$, and a complementary price λ. There are at most two positive basis columns and their total allocation is below $m/8$. The basis and price stay fixed until exit.

Let O be the fully determined outside count. After some rounds let r roots per type remain untouched, and let $\mathscr{F}_r$ be the sigma-field of all responses and random choices so far. All processed roots have

known actions; define $L^{\mathrm{ret}} = O + \sum_{g \text{ processed}} l_g$ and $U^{\mathrm{ret}} = O + \sum_{g \text{ processed}} u_g$. For $p_j = \mathbb{E}X_j$ the actual remaining capacities are

$$c_r^+ = L^{\mathrm{ret}} + r\sum_j p_j - a, \qquad c_r^- = b - U^{\mathrm{ret}} - r\sum_j p_j. \tag{79}$$

Omit nonexistent outer sides. A round processes one fresh root of each type. Its observed endpoints satisfy

$$\Delta c_r = c_{r-1} - c_r = \left(\sum_j (l_j - p_j), \sum_j (p_j - u_j)\right), \qquad \|\Delta c_r\|_\infty \le J x_{\max}. \tag{80}$$

For two columns put $\mathbf{Q} = [B_{\pi_1}, B_{\pi_2}]$, $z_r = \mathbf{Q}^{-1} c_r$. For one column binding side s, put $z_r = c_r^s / B_\pi^s$ and, if the other side exists, $e_r = c_r^{-s} - B_\pi^{-s} z_r$. If both sides are tight at entry, apply the exit allocation immediately; its entry distance to a knot is zero. Otherwise the other side is absent or strictly slack at entry, so complementarity gives $\lambda_{-s} = 0$. In every case $G_{\pi_i} = \lambda \cdot B_{\pi_i}$. Define a raw-increment bound using the selected catalog and its nonsingular two-column bases:

$$A = 1 + J x_{\max} \max\left\{\kappa_{\mathrm{bas}}, \max_{\pi \in \mathcal{C}, s} \frac{1}{B_\pi^s}, \max_{\pi \in \mathcal{C}, s} \left(1 + \frac{B_\pi^{-s}}{B_\pi^s}\right)\right\}.$$

Set the first term to zero when every basis has one column. For the three possible coordinates,

$$\begin{aligned} \|\Delta z_r\|_\infty &= \|\mathbf{Q}^{-1}\Delta c_r\|_\infty \le \|\mathbf{Q}^{-1}\|_\infty J x_{\max} \le A, \\ |\Delta z_r| &= |\Delta c_r^s| / B_\pi^s \le J x_{\max} / B_\pi^s \le A, \\ |\Delta e_r| &= |\Delta c_r^{-s} - (B_\pi^{-s}/B_\pi^s)\Delta c_r^s| \le J x_{\max}(1 + B_\pi^{-s}/B_\pi^s) \le A. \end{aligned}$$

Feasibility and complementary slackness then give

$$|\Delta z_{r,i}| \le A, \quad |\Delta e_r| \le A, \qquad z_r \ge 0,\ e_r \ge 0 \implies \mathcal{V}(c_r) = \lambda \cdot c_r. \tag{81}$$

The bound includes the finite basis inverse norms in (45); it imposes no uniqueness of an optimum.

One activation per coefficient. Fix $\varepsilon = 1/32$ and an integer $r_0 \ge \max\{2C_* + 2, 128(A+1)\}$. At entry mark coefficient i active if $z_{m,i}/m \ge \varepsilon$, and dormant otherwise. Before each round stop if $r \le r_0$, any $z_{r,i} \le 0$, any $z_{r,i}/r \ge 1/4$, or an existing nonbinding slack is nonpositive. Otherwise activate every dormant coefficient whose ratio is at least ε, and set

$$q_{r,i} = \begin{cases} z_{r,i}/r, & i \text{ has been activated}, \\ 0, & i \text{ is dormant}. \end{cases} \tag{82}$$

Activation is permanent. Select each partial policy with this probability on its type; use one categorical choice for columns of the same type, and independent choices across types. All other roots use minimizing full policies. The total probability is below $1/2$, so every assignment is feasible. Execute and pay every query on the selected path, then update actual endpoints. A zero entry coefficient is removed from the basis; a zero entry slack invokes exit correction immediately.

Let $q_{r,j}^{\mathrm{full}} = 1 - \sum_{i:\, j(i)=j} q_{r,i}$. A full policy has expected endpoints p_j in both directions. Thus

$$\begin{aligned} \mathbb{E}[l_j - p_j \mid \mathscr{F}_r] &= q_{r,j}^{\mathrm{full}}(p_j - p_j) + \sum_{i:\, j(i)=j} q_{r,i}(\mathbb{E}l_{\pi_i} - p_j) = -\sum_{i:\, j(i)=j} q_{r,i} B_{\pi_i}^+, \\ \mathbb{E}[p_j - u_j \mid \mathscr{F}_r] &= q_{r,j}^{\mathrm{full}}(p_j - p_j) + \sum_{i:\, j(i)=j} q_{r,i}(p_j - \mathbb{E}u_{\pi_i}) = -\sum_{i:\, j(i)=j} q_{r,i} B_{\pi_i}^-. \end{aligned}$$

Sum over types and apply the fixed linear maps defining z_r, e_r:

$$\begin{aligned} \mathbb{E}[\Delta c_r \mid \mathscr{F}_r] &= -\sum_i B_{\pi_i} q_{r,i}, & \mathbb{E}[\Delta z_{r,i} \mid \mathscr{F}_r] &= -q_{r,i}, \\ \mathbb{E}[\Delta e_r \mid \mathscr{F}_r] &= 0, & \mathbb{E}[\lambda \cdot (c_r - c_{r-1}) \mid \mathscr{F}_r] &= \sum_i G_{\pi_i} q_{r,i}. \end{aligned} \tag{83}$$

For a dormant coefficient $q_{r,i} = 0$, so $\mathbb{E}[z_{r-1,i} \mid \mathscr{F}_r] = z_{r,i}$. After activation, $q_{r,i} = z_{r,i}/r$ and

$$\mathbb{E}\left[\frac{z_{r-1,i}}{r-1}\,\middle|\,\mathscr{F}_r\right] = \frac{z_{r,i} - q_{r,i}}{r-1} = \frac{z_{r,i} - z_{r,i}/r}{r-1} = \frac{z_{r,i}}{r},$$
$$\frac{z_{r-1,i}}{r-1} - \frac{z_{r,i}}{r} = \frac{\Delta z_{r,i} + z_{r,i}/r}{r-1}.$$

Before exit $z_{r,i}/r < 1/4$. The raw increment bound gives

$$\left|\frac{z_{r-1,i}}{r-1} - \frac{z_{r,i}}{r}\right| \le \frac{A+1}{r-1}. \tag{84}$$

An initial active ratio is in $[\varepsilon, 1/8]$. At a later activation,

$$\varepsilon \le z_{r,i}/r \le \varepsilon + (A+\varepsilon)/r < 1/8.$$

Thus every active coordinate starts a fixed distance from both exit boundaries, including a coordinate that was arbitrarily small at entry.

Fixed-price accounting through exit. Let R be the number of fresh roots of each type at exit, and let $d_b \le 2$ be the number of basis columns. The increment bound and the last pre-exit ratios imply, for every coefficient and existing nonbinding slack,

$$-A \le z_{R,i} \le (R+1)/4 + A, \qquad e_R \ge -A, \qquad \sum_i (z_{R,i})_+ \le R/2 + 2A + 1/2.$$

These inequalities also hold at an immediate entry exit: its coefficients are nonnegative and at most $R/8$. Put

$$R_{\mathrm{cut}} = \lceil 4A + 2 \rceil, \quad G_{\max} = \max_i G_{\pi_i}, \quad \mathscr{E}_{\mathrm{exit}} = A(2x_{\max} + 1),$$
$$c_{\mathrm{fee}} = G_{\max} \max\{2, (R_{\mathrm{cut}} + 1)/2 + 2A\}.$$

If $R \ge R_{\mathrm{cut}}$, set $n_i = \lfloor (z_{R,i})_+ \rfloor$ and assign that many fresh roots to policy π_i; otherwise set all $n_i = 0$. All unassigned roots are fully evaluated. For large R, $\sum_i n_i \le R$, including when both columns use the same type, and

$$-A \le z_{R,i} - n_i < 1, \qquad \lambda \cdot c_R = \sum_i G_{\pi_i} z_{R,i}, \qquad \lambda \cdot c_R - \sum_i G_{\pi_i} n_i \le 2G_{\max}.$$

For small R, the coefficient bound gives $\lambda \cdot c_R \le G_{\max}[(R_{\mathrm{cut}} + 1)/2 + 2A]$. Consequently in both cases

$$\sum_i G_{\pi_i} n_i \ge \lambda \cdot c_R - c_{\mathrm{fee}}. \tag{85}$$

The capacity identity is $c_R = \sum_i B_{\pi_i} z_{R,i}$ when both sides bind. With one binding side, add e_R in the other coordinate; that coordinate has zero price by complementarity. Thus, for each existing side,

$$c_R^s - \sum_i B_{\pi_i}^s n_i \ge -2x_{\max} A - A = -\mathscr{E}_{\mathrm{exit}}.$$

For small R, use $n_i = 0$ and $z_{R,i} \ge -A$ to obtain the same inequality. Let $[L_1, U_1]$ include the fixed earlier endpoints and all just-processed fresh roots. Conditional on $\mathscr{F}_R$, the assignments were chosen before their independent responses. By the net-column identities,

$$\mathbb{E}(L_1 - a \mid \mathscr{F}_R) = c_R^+ - \sum_i B_{\pi_i}^+ n_i \ge -\mathscr{E}_{\mathrm{exit}},$$
$$\mathbb{E}(b - U_1 \mid \mathscr{F}_R) = c_R^- - \sum_i B_{\pi_i}^- n_i \ge -\mathscr{E}_{\mathrm{exit}},$$
$$\mathrm{Var}(L_1 \mid \mathscr{F}_R), \quad \mathrm{Var}(U_1 \mid \mathscr{F}_R) \le JR{x_{\max}}^2/4.$$

For a real variable Z of mean μ_Z and variance v_Z,

$$\mathbb{E}(-Z)_+ = \tfrac{1}{2}(\mathbb{E}|Z| - \mu_Z) \le \tfrac{1}{2}(\sqrt{\mu_Z^2 + v_Z} - \mu_Z) \le (-\mu_Z)_+ + \tfrac{1}{2}\sqrt{v_Z}.$$

Let c_{rec} bound both completion prices over the fixed core histories. All retained roots use those histories. With $\mathcal{D}_1 = (a - L_1)_+ + (U_1 - b)_+$,

$$\begin{aligned}
\mathbb{E}[\mathcal{D}_1 \mid \mathscr{F}_R] &\le 2\mathscr{E}_{\mathrm{exit}} + \tfrac{1}{2}\sqrt{JRx_{\max}^2/4} + \tfrac{1}{2}\sqrt{JRx_{\max}^2/4} \\
&= 2\mathscr{E}_{\mathrm{exit}} + (x_{\max}/2)\sqrt{JR}, \\
\mathbb{E}[C_{\mathrm{repair}} \mid \mathscr{F}_R] &\overset{\text{Lemma } B}{\le} c_{\mathrm{rec}}\{\mathbb{E}[\mathcal{D}_1 \mid \mathscr{F}_R] + 2x_{\max}\} \\
&\le 2c_{\mathrm{rec}}(\mathscr{E}_{\mathrm{exit}} + x_{\max}) + (c_{\mathrm{rec}}x_{\max}/2)\sqrt{JR}, \\
\mathbb{E}[C_{\mathrm{exit}} \mid \mathscr{F}_R] &= R\sum_j v_j - \sum_i G_{\pi_i} n_i + \mathbb{E}[C_{\mathrm{repair}} \mid \mathscr{F}_R] \\
&\le R\sum_j v_j - \lambda \cdot c_R + c_{\mathrm{fee}} + 2c_{\mathrm{rec}}(\mathscr{E}_{\mathrm{exit}} + x_{\max}) + (c_{\mathrm{rec}}x_{\max}/2)\sqrt{JR}.
\end{aligned}$$

Define

$$A_1 = \max\{c_{\mathrm{fee}}, 2\mathscr{E}_{\mathrm{exit}}\}, \qquad A_2 = c_{\mathrm{fee}} + 2c_{\mathrm{rec}}(\mathscr{E}_{\mathrm{exit}} + x_{\max}) + c_{\mathrm{rec}}x_{\max}\sqrt{J}/2.$$

Combining these bounds gives

$$\begin{aligned}
\sum_i G_{\pi_i} n_i &\ge \lambda \cdot c_R - A_1, \\
\mathbb{E}[(a - L_1)_+ + (U_1 - b)_+ \mid \mathscr{F}_R] &\le A_1 + x_{\max}\sqrt{JR}, \\
\mathbb{E}[C_{\mathrm{exit}} \mid \mathscr{F}_R] &\le R\sum_j v_j - \lambda \cdot c_R + A_2(1 + \sqrt{R}).
\end{aligned} \tag{86}$$

Absent sides contribute zero. A positive nonbinding slack causes no deficit and has zero dual price, regardless of its size.

Since a full round costs $\sum_j v_j$ in expectation before partial-policy gains,

$$\sum_j v_j - \mathbb{E}[C_{\mathrm{round}} \mid \mathscr{F}_r] = \sum_i G_{\pi_i} q_{r,i} = \mathbb{E}[\lambda \cdot (c_r - c_{r-1}) \mid \mathscr{F}_r]. \tag{87}$$

Let $H_r = \mathbf{1}_{\{r>R\}}$ indicate an executed tracking round. It is measurable before that round. The fixed price is measurable at entry. Hence

$$\begin{aligned}
\mathbb{E}\left[\sum_{r=R+1}^{m} C_r \,\middle|\, \mathscr{F}\right] &= \mathbb{E}\left[(m - R)\sum_j v_j - \sum_{r=1}^{m} H_r \lambda \cdot (c_r - c_{r-1}) \,\middle|\, \mathscr{F}\right] \\
&= \mathbb{E}\left[(m - R)\sum_j v_j - \lambda \cdot (c_m - c_R) \,\middle|\, \mathscr{F}\right], \\
\mathbb{E}[C_{\mathrm{remaining}} \mid \mathscr{F}] &\overset{(86)}{\le} m\sum_j v_j - \lambda \cdot c_m + A_2(1 + \mathbb{E}[\sqrt{R} \mid \mathscr{F}]).
\end{aligned}$$

Since $\lambda \cdot c_m = \mathcal{F}^{\mathrm{net}}(M)$, this is

$$\mathbb{E}[C_{\mathrm{remaining}} \mid \mathscr{F}] \le m\sum_j v_j - \mathcal{F}^{\mathrm{net}}(M) + A_2(1 + \mathbb{E}[\sqrt{R} \mid \mathscr{F}]). \tag{88}$$

Full evaluations made while a coefficient is dormant remain in this equality. Prices and activations are measurable in the purchased transcript.

Dormant and active exit probabilities. Let $\mathcal{K}_N$ contain genuine cell endpoints and the finitely many internal ray intersections

$$\xi_\pi = \frac{B_\pi^- a + B_\pi^+ b}{B_\pi^+ + B_\pi^-}, \qquad d_{\mathrm{knot}} = \mathrm{dist}(M, \mathcal{K}_N).$$

The basis slopes in (45) imply that every entry coefficient, and every existing nonbinding slack, is at least $a_0 d_{\mathrm{knot}}$ for a fixed $a_0 > 0$. A dormant coefficient has zero raw drift until activation or exit, with at most m increments bounded by A. The same holds for a nonbinding slack. For a stopped raw coordinate Z_t, put $h = a_0 d_{\mathrm{knot}}$. Conditional Hoeffding and the maximal supermartingale inequality give

$$\begin{aligned}\mathbb{E}[e^{-\theta(Z_{t+1}-Z_t)} \mid \mathscr{F}_t] &\le e^{\theta^2 A^2/2},\\ \Pr\{\min_{t\le m}(Z_t - Z_0) \le -h \mid \mathscr{F}\} &\le \inf_{\theta>0} e^{-\theta h + m\theta^2 A^2/2}\\ &= \exp\left\{-\frac{h^2}{mA^2} + \frac{mA^2}{2}\left(\frac{h}{mA^2}\right)^2\right\} = e^{-h^2/(2mA^2)}.\end{aligned}$$

For $h = 0$ use the probability bound 1. Summing over at most three coordinates gives

$$\Pr(\text{dormant or slack exit} \mid \mathscr{F}) \le 3\exp[-a_0^2 {d_{\mathrm{knot}}}^2/(2mA^2)]. \tag{89}$$

For an active coordinate, condition on its actual activation history. The normalized starting value lies in $[\varepsilon, 1/8]$, so its distance from either endpoint of $(0, 1/4)$ is at least ε. For an activation at remaining size r_a,

$$\sum_{r=t+1}^{r_a} \frac{(A+1)^2}{(r-1)^2} \le (A+1)^2 \sum_{j=t}^{\infty} \frac{1}{j(j-1)} = (A+1)^2 \sum_{j=t}^{\infty}\left(\frac{1}{j-1} - \frac{1}{j}\right) = \frac{(A+1)^2}{t-1}.$$

Apply the same exponential inequality to the normalized martingale:

$$\begin{aligned}\Pr\left(\max_{t\le r\le r_a}\left|\frac{z_{r,i}}{r} - \frac{z_{r_a,i}}{r_a}\right| \ge \varepsilon \,\middle|\, \mathscr{F}_{r_a}\right) &\le 2\inf_{\theta>0}\exp\left[-\theta\varepsilon + \frac{\theta^2}{2}\sum_{r=t+1}^{r_a}\frac{(A+1)^2}{(r-1)^2}\right]\\ &\le 2\exp[-\varepsilon^2(t-1)/(2(A+1)^2)].\end{aligned}$$

If $\mathcal{A}_i$ is the event that coefficient i activates, then

$$\Pr(R \ge t,\ \text{active exit} \mid \mathscr{F}) \le \sum_{i=1}^{d_b} \mathbb{E}[\mathbf{1}_{\mathcal{A}_i} 2e^{-\varepsilon^2(t-1)/(2(A+1)^2)} \mid \mathscr{F}] \le 4e^{-\varepsilon^2(t-1)/(2(A+1)^2)}.$$

Thus

$$\Pr(R \ge t,\ \text{active-coordinate exit} \mid \mathscr{F}) \le 4\exp\left[-\frac{\varepsilon^2(t-1)}{2(A+1)^2}\right] \quad (t > r_0). \tag{90}$$

The activation times need not be deterministic: the bound holds conditionally on each activation history, uniformly in the remaining horizon, and the tower property removes that conditioning. A dormant ratio crossing $1/4$ must either activate first or occur at bounded r, since its first ε crossing is below $1/8$. Thus all exits have been counted. Set $a_{\mathrm{act}} = \varepsilon^2/[2(A+1)^2]$ and $a_{\mathrm{raw}} = a_0^2/(2A^2)$. Integrating the two tails gives

$$\begin{aligned}\mathbb{E}[\sqrt{R} \mid \mathscr{F}] &= \int_0^m \frac{\Pr(R > t \mid \mathscr{F})}{2\sqrt{t}}\,dt\\ &\le \int_0^{r_0}\frac{dt}{2\sqrt{t}} + \int_{r_0}^{\infty}\frac{4e^{-a_{\mathrm{act}}(t-1)}}{2\sqrt{t}}\,dt + 3e^{-a_{\mathrm{raw}}d_{\mathrm{knot}}^2/m}\int_0^m\frac{dt}{2\sqrt{t}}\\ &\le \sqrt{r_0} + 2e^{a_{\mathrm{act}}}\frac{\Gamma(1/2)}{\sqrt{a_{\mathrm{act}}}} + 3\sqrt{m}\,e^{-a_{\mathrm{raw}}d_{\mathrm{knot}}^2/m}\\ &= \sqrt{r_0} + 2e^{a_{\mathrm{act}}}\sqrt{\pi/a_{\mathrm{act}}} + 3\sqrt{m}\,e^{-a_{\mathrm{raw}}d_{\mathrm{knot}}^2/m}.\end{aligned}$$

Consequently, with $A_5 = 3 + \sqrt{r_0} + 2e^{a_{\mathrm{act}}}\sqrt{\pi/a_{\mathrm{act}}}$ and $a_5 = a_{\mathrm{raw}}$,

$$\mathbb{E}[\sqrt{R} \mid \mathscr{F}] \le A_5[1 + \sqrt{m}\exp(-a_5 {d_{\mathrm{knot}}}^2/m)]. \tag{91}$$

Delayed activation removes the reciprocal-distance exit term: small coefficients fluctuate on their raw scale until their probability is bounded away from zero. The cost of that delay is already included in (87).

Averaging the conditional bound. Given $\mathscr{F}$, $K - M$ is a sum of Jm independent centered bounded variables. Let s_M be the slope of the affine piece containing M, choosing either adjacent slope at a knot. For every real x,

$$|\mathcal{F}^{\mathrm{net}}(x) - \mathcal{F}^{\mathrm{net}}(M) - s_M(x - M)| \le 2L_{\mathrm{net}}(|x - M| - d_{\mathrm{knot}})_+.$$

Before the first knot the left side is zero; afterwards its derivative in absolute value is at most $2L_{\mathrm{net}}$. Since $\mathbb{E}[K - M \mid \mathscr{F}] = 0$, conditional Hoeffding and integration give

$$\begin{aligned}
|\mathbb{E}[\mathcal{F}^{\mathrm{net}}(K) \mid \mathscr{F}] - \mathcal{F}^{\mathrm{net}}(M)| &\le 4L_{\mathrm{net}} \int_{d_{\mathrm{knot}}}^{\infty} e^{-2t^2/(Jm{x_{\max}}^2)}\, dt \\
&= 4L_{\mathrm{net}} \int_0^{\infty} e^{-2(d_{\mathrm{knot}}+s)^2/(Jmx_{\max}^2)}\, ds \\
&\le 4L_{\mathrm{net}} e^{-2d_{\mathrm{knot}}^2/(Jmx_{\max}^2)} \int_0^{\infty} e^{-2s^2/(Jmx_{\max}^2)}\, ds \\
&= \sqrt{2\pi} L_{\mathrm{net}} x_{\max} \sqrt{Jm} e^{-2d_{\mathrm{knot}}^2/(Jmx_{\max}^2)} \\
&\le \sqrt{2\pi} L_{\mathrm{net}} x_{\max} \sqrt{Jm} e^{-d_{\mathrm{knot}}^2/(Jmx_{\max}^2)}.
\end{aligned}$$

Set $A_6 = \sqrt{2\pi} L_{\mathrm{net}} x_{\max} \sqrt{J}$ and $a_6 = (J{x_{\max}}^2)^{-1}$,

$$|\mathbb{E}[\mathcal{F}^{\mathrm{net}}(K) \mid \mathscr{F}] - \mathcal{F}^{\mathrm{net}}(M)| \le A_6 \sqrt{m} \exp(-a_6 {d_{\mathrm{knot}}}^2/m). \tag{92}$$

Set $c_{\mathrm{tr}} = \min(a_5, a_6)$. The three estimates combine in the following order:

$$\begin{aligned}
\mathbb{E}[C_{\mathrm{remaining}} \mid \mathscr{F}] &\overset{(88)}{\le} m \sum_j v_j - \mathcal{F}^{\mathrm{net}}(M) + A_2(1 + \mathbb{E}[\sqrt{R} \mid \mathscr{F}]) \\
&\overset{(91)}{\le} m \sum_j v_j - \mathcal{F}^{\mathrm{net}}(M) + A_2(1 + A_5) + A_2 A_5 \sqrt{m}\, e^{-a_5 d_{\mathrm{knot}}^2/m} \\
&\overset{(92)}{\le} m \sum_j v_j - \mathbb{E}[\mathcal{F}^{\mathrm{net}}(K) \mid \mathscr{F}] + A(1 + \sqrt{m}\, e^{-c_{\mathrm{tr}} d_{\mathrm{knot}}^2/m}),
\end{aligned}$$

where $A = A_2(1 + A_5) + A_6$. For the averaged statement below, intersect every event with $\mathcal{E}$; no conditioning on $\mathcal{E}$ is applied to the marginal mass bound for K.

Fix a nearest knot ξ, write $\ell_\xi = d_\Gamma(\xi)$ and $\ell_{\xi,+} = \ell_\xi + 1$. On $|M - \xi| \le u$, (76) implies $m \le c_*(\ell_\xi + u) + C_* =: m_{\mathrm{cap}}(u)$. Put $E_{\xi,u} = \{\mathcal{E}, |M - \xi| \le u, \xi \text{ nearest}\}$. Conditional independence and Chebyshev give

$$\begin{aligned}
\mathrm{Var}(K \mid \mathscr{F}) &\le Jmx_{\max}^2/4, \\
\Pr(|K - M| > x_{\max}\sqrt{Jm} \mid \mathscr{F}) &\le 1/4, \\
\tfrac{3}{4} \Pr(E_{\xi,u}) &\le \mathbb{E}[\mathbf{1}_{E_{\xi,u}} \Pr(|K - M| \le x_{\max}\sqrt{Jm} \mid \mathscr{F})] \\
&\le \Pr(|K - \xi| \le u + x_{\max}\sqrt{Jm_{\mathrm{cap}}(u)}).
\end{aligned}$$

An interval of radius v contains at most $2v + 1$ integers. Therefore

$$\begin{aligned}
\Pr\{\mathcal{E},\ |M - \xi| \le u,\ \xi \text{ nearest}\} &\le \tfrac{4}{3} \Pr\{|K - \xi| \le u + x_{\max}\sqrt{Jm_{\mathrm{cap}}(u)}\} \\
&\le A_7 \mathfrak{p}_N[u + \sqrt{\ell_\xi + u + 1} + 1].
\end{aligned} \tag{93}$$

The bound uses the lattice masses of K through the conditional law at M. For $d_{\mathrm{knot}} \le \ell_{\xi,+}$, one has $m \le A_8 \ell_{\xi,+}$, so

$$\sqrt{m} e^{-c_{\mathrm{tr}} {d_{\mathrm{knot}}}^2/m} \le A_9 \sqrt{\ell_{\xi,+}} e^{-a_9 {d_{\mathrm{knot}}}^2/\ell_{\xi,+}}.$$

Let $F_\xi(u) = \Pr(\mathcal{E}, d_{\mathrm{knot}} \le u, \xi \text{ nearest})$. Integration by parts and (93) give

$$\begin{aligned}
&\mathbb{E}[e^{-a_9 {d_{\mathrm{knot}}}^2/\ell_{\xi,+}} \mathbf{1}\{\mathcal{E}, d_{\mathrm{knot}} \le \ell_{\xi,+}, \xi \text{ nearest}\}] \\
&= e^{-a_9 \ell_{\xi,+}} F_\xi(\ell_{\xi,+}) + \int_0^{\ell_{\xi,+}} \frac{2a_9 u}{\ell_{\xi,+}} e^{-a_9 u^2/\ell_{\xi,+}} F_\xi(u)\, du \\
&\le A_{10} \mathfrak{p}_N \sqrt{\ell_{\xi,+}}.
\end{aligned} \tag{94}$$

To justify the last inequality, for $0 \le u \le \ell_{\xi,+}$ and $\ell_{\xi,+} = \ell_\xi + 1 \ge 1$, (93) is at most $A_7 \mathfrak{p}_N (u + \sqrt{2\ell_{\xi,+}} + 1)$. In its integral, the terms u and $\sqrt{2\ell_{\xi,+}} + 1$ are bounded respectively using

$$\int_0^\infty \frac{2a_9 u^2}{\ell_{\xi,+}} e^{-a_9 u^2/\ell_{\xi,+}}\, du = \frac{\sqrt{\pi}}{2\sqrt{a_9}} \sqrt{\ell_{\xi,+}}, \qquad \int_0^\infty \frac{2a_9 u}{\ell_{\xi,+}} e^{-a_9 u^2/\ell_{\xi,+}}\, du = 1.$$

The endpoint term satisfies $e^{-a_9 \ell_{\xi,+}} (\ell_{\xi,+} + \sqrt{2\ell_{\xi,+}} + 1) \le C(a_9) \sqrt{\ell_{\xi,+}}$ by maximizing $\sqrt{\ell_{\xi,+}} e^{-a_9 \ell_{\xi,+}}$ over $\ell_{\xi,+} \ge 1$. All constants in (94) are thus independent of N, ℓ_ξ. For $d_{\mathrm{knot}} > \ell_{\xi,+}$, $m \le A_{11} d_{\mathrm{knot}}$ and $\sqrt{m} e^{-c_{\mathrm{tr}} {d_{\mathrm{knot}}}^2/m} \le \sqrt{A_{11} d_{\mathrm{knot}}}\, e^{-c_{\mathrm{tr}} d_{\mathrm{knot}}/A_{11}}$ is bounded by one constant on the union of these events: every entry has one nearest knot. Summing only the near-knot contributions therefore gives

$$\mathbb{E}[\mathbf{1}_{\mathcal{E}} \sqrt{m} e^{-c_{\mathrm{tr}} {d_{\mathrm{knot}}}^2/m}] \le A_{12} \left[1 + \mathfrak{p}_N \sum_{\xi \in \mathcal{K}_N} (d_\Gamma(\xi) + 1) \right]. \tag{95}$$

This proves (78) for any finite number of thresholds. All constants use only $J, x_{\max}, \beta_{\mathrm{gen}}$, the selected basis inverse bounds, and the fees and completion prices of executed policies. They remain uniform when laws or graphs vary while these quantities stay bounded and the indicated positive bounds stay away from zero. In fact the two-column slope formula gives $|z_i'| \ge 1/x_{\max}$, because $|\det Q| \le x_{\max}(B_i^+ + B_i^-)$ for the opposite column; one-column coefficient and slack slopes obey the same lower bound. Thus no additional vanishing-slope hypothesis is needed. Initial supply is handled in Lemma D. For the stated width bound choose a nearest knot ξ and put $d = |M - \xi|$. Since $d_\Gamma(\xi) \le w_\Gamma$, $m \le c_*(w_\Gamma + d) + C_*$. If $d \le w_\Gamma + 1$, the kernel is at most $\sqrt{(2c_* + C_*)(w_\Gamma + 1)}$. If $d > w_\Gamma + 1$, put $A_0 = 2c_* + C_*$ and use

$$\sqrt{m} e^{-c_{\mathrm{tr}} d^2/m} \le \sqrt{A_0 d}\, e^{-c_{\mathrm{tr}} d/A_0} \le A_0/\sqrt{2 e c_{\mathrm{tr}}}.$$

This is a pointwise bound on every entry transcript, so it introduces neither a variance condition nor a factor counting the knots. □

A normal approximation to the expected saving. Extend $F_N = \mathcal{F}^{\mathrm{net}}$ affinely beyond the two outer endpoints; it is globally L_{net}-Lipschitz. Put $\mu_N = \mathbb{E}K$, $\sigma_N^2 = \mathrm{Var}\, K$ and $T_3 = \sum_g \mathbb{E}|X_g - \mathbb{E}X_g|^3$. If $\sigma_N = 0$, then $K = \mu_N$ almost surely and its expected envelope is $F_N(\mu_N)$. For independent centered Y_g with finite third absolute moments and $\sum_g \mathbb{E}Y_g^2 = \sigma^2 > 0$, put $\xi_g = Y_g/\sigma$, $v_g = \mathbb{E}\xi_g^2$ and $W = \sum_g \xi_g$. Omit terms with $v_g = 0$. Let ξ_g^z have the zero-bias law of ξ_g, defined by $\mathbb{E}[\xi_g f(\xi_g)] = v_g \mathbb{E} f'(\xi_g^z)$. Choose these variables independently of all original summands and one another. Independently choose I with $\Pr(I = g) = v_g$. The zero-bias construction and normal bound (Ross, 2011, Theorem 3.29 and Propositions 3.30, 3.32) give, for $Z \sim N(0,1)$ and $f(x) = x|x|/2$,

$$\begin{aligned}
W^z &= W - \xi_I + \xi_I^z, \\
v_g \mathbb{E}|\xi_g^z| &= v_g \mathbb{E} f'(\xi_g^z) = \mathbb{E}[\xi_g f(\xi_g)] = \tfrac{1}{2}\mathbb{E}|\xi_g|^3, \\
v_g \mathbb{E}|\xi_g| &\overset{\text{Cauchy–Schwarz}}{\le} v_g^{3/2} \overset{\text{Hölder}}{\le} \mathbb{E}|\xi_g|^3, \\
d_W(W, Z) &\le 2\mathbb{E}|W^z - W| \\
&\le 2 \sum_g v_g \big(\mathbb{E}|\xi_g| + \mathbb{E}|\xi_g^z| \big) \\
&\le 3 \sum_g \mathbb{E}|\xi_g|^3 = \frac{3 \sum_g \mathbb{E}|Y_g|^3}{\sigma^3}.
\end{aligned}$$

Thus

$$d_W\left(\frac{\sum_g Y_g}{\sigma}, Z \right) \le \frac{3 \sum_g \mathbb{E}|Y_g|^3}{\sigma^3}, \qquad d_W(U, V) = \sup_{\mathrm{Lip}(f) \le 1} |\mathbb{E}f(U) - \mathbb{E}f(V)|.$$

If $L_{\mathrm{net}} = 0$, then F_N is constant and the approximation error is zero. Otherwise use $Y_g = X_g - \mathbb{E}X_g$, $\sigma = \sigma_N$ and $f(z) = F_N(\mu_N + \sigma_N z)/(L_{\mathrm{net}}\sigma_N)$. This f is one-Lipschitz. Therefore

$$\begin{aligned}|\mathbb{E}F_N(K) - \mathbb{E}F_N(\mu_N + \sigma_N Z)| &\le L_{\mathrm{net}}\sigma_N d_W((K-\mu_N)/\sigma_N, Z)\\ &\le 3L_{\mathrm{net}}T_3/\sigma_N^2\\ &\le 3L_{\mathrm{net}}x_{\max}\frac{\sum_g \mathbb{E}(X_g - \mathbb{E}X_g)^2}{\sigma_N^2} = 3L_{\mathrm{net}}x_{\max}.\end{aligned}$$

Equivalently,

$$|\mathbb{E}F_N(K) - \mathbb{E}F_N(\mu_N + \sigma_N Z)| \le \frac{3L_{\mathrm{net}}T_3}{\sigma_N^2} \le 3L_{\mathrm{net}}x_{\max}, \qquad Z \sim N(0,1). \tag{96}$$

Once its local prices are computed, write $F_N(x) = \alpha_0 + \alpha_1 x + \sum_j \Delta_j (x - t_j)_+$ using its finitely many slope changes. Then, with $z_j = (t_j - \mu_N)/\sigma_N$,

$$\mathbb{E}F_N(\mu_N + \sigma_N Z) = \alpha_0 + \alpha_1\mu_N + \sum_j \Delta_j \left[\sigma_N\phi(z_j) + (\mu_N - t_j)(1 - \Phi(z_j))\right]. \tag{97}$$

Thus a one-dimensional normal integral preserves the bounded remainder. The local response model determines the net curve; the normal integral then evaluates its expected value from the count moments.

C.3 Library supply and the count scale

Let $\mathcal{T}_0$ be a fixed finite family of local response types and let $\mathcal{D}_0$ be its net-price region from (31). Write $\mathcal{P}_0$ for the finite vertices of $\mathcal{D}_0$. The family *determines the prices* of the batch if

$$U_{g,\lambda}(\mathfrak{h}_g^0) = v_g(\mathfrak{h}_g^0) \quad (g = 1, \dots, N,\ \lambda \in \mathcal{P}_0). \tag{98}$$

Every vertex is checked by a local backward calculation. Nonnegative rays and $B_{g\pi} \ge 0$ extend this inequality from vertices to all of $\mathcal{D}_0$:

$$(98) \iff G_{g\pi} \le \lambda \cdot B_{g\pi} \quad (\lambda \in \mathcal{D}_0,\ g, \pi) \iff \mathcal{D}_{\mathrm{batch}} = \mathcal{D}_0. \tag{99}$$

The last equivalence uses the presence of the core types in the batch. Consequently the batch envelope is exactly the fixed core envelope, while V_0 includes every component's actual full-evaluation value.

Lemma D (Finite library supply)**.** *Assume (P), bounded integer contributions, and any finite nonempty threshold family. Select $m_0 \ge 1$ actual copies of each of J_0 local types, and a nonempty positive-gain policy library with $\beta_{\mathrm{gen}} > 0$. Its envelope is $F_0 = \mathcal{F}_0^{\mathrm{net}}$, its axial-price sum is Λ_0, and its knot statistic is $\mathfrak{K}_N$ from Lemma C. Put*

$$\mu_N = \mathbb{E}K, \quad \sigma_N^2 = \operatorname{Var} K, \quad d_0 = d_\Gamma(\mu_N), \quad \tau_N = 1 + \sigma_N + d_0,$$
$$c_0 = 8/\beta_{\mathrm{gen}}, \quad c_1 = 1 + c_0 J_0 x_{\max}.$$

Complete the exterior and let M_0 be the resulting posterior mean. For $\mathcal{E}_0 = \{m_0 \le c_0 d_\Gamma(M_0) + c_1\}$ define $\mathcal{L}_N(m_0) = \mathbb{E}[F_0(K)\mathbf{1}_{\mathcal{E}_0}]$. There is a support-exact policy satisfying

$$\mathbb{E}C_{\mathsf{Track}_0} \le V_0 - \mathbb{E}F_0(K) + \mathcal{L}_N(m_0) + A(1 + \mathfrak{p}_N\mathfrak{K}_N). \tag{100}$$

Its constant uses only the executed library geometry and recourse quantities listed after (95)*. No background price equality, background support anchor, or variance lower bound is required. For $\kappa \in \{1, 2\}$ set*

$$\delta_\kappa(t) = \begin{cases} \min\{1, 2\exp[-t^2/(2\kappa\sigma_N^2 + 2\kappa x_{\max}t/3)]\}, & t > 0,\\ 1, & t \le 0.\end{cases} \tag{101}$$

Then, with $t = (m_0 - c_1)/c_0 - d_0$,

$$\mathcal{L}_N(m_0) \le \sqrt{2}\Lambda_0\sqrt{\sigma_N^2 + d_0^2}\sqrt{\delta_1(t)}. \tag{102}$$

In particular, the sufficient supply

$$m_0 \ge c_0\left[d_0 + 2\sigma_N\sqrt{\log(2+\tau_N)} + \tfrac{4}{3}x_{\max}\log(2+\tau_N)\right] + c_1 \tag{103}$$

makes $\mathcal{L}_N(m_0) \le 2\Lambda_0$. Under the original (K), fixed library conditions, and

$$m_0 = \lceil L_{\mathrm{core}}\sqrt{N\log(2N)}\rceil, \qquad L_{\mathrm{core}} \ge 16x_{\max}/\beta_{\mathrm{gen}}, \tag{104}$$

the upper bound becomes $V_0 - \mathbb{E}F_0(K) + O(1)$.

Proof. Core entry and exceptional supply. Choose m_0 roots of each core type before observing any response. Fully evaluate all other components using their minimizing routines. Let $\mathscr{F}_0$ be the sigma-field generated by the resulting transcript and $M_0 = \mathbb{E}[K \mid \mathscr{F}_0]$. The reserved roots are still independent with their specified core laws. An empty positive core is handled by full evaluation, whose expected fee is V_0. Otherwise put $\beta_0 = \beta_{\mathrm{gen}} = \min_{\pi\in\mathcal{C},s} B^s_\pi > 0$, $c_0 = 8/\beta_0$, $c_1 = 1 + c_0 J_0 x_{\max}$, and let

$$\mathcal{E}_0 = \{m_0 \le c_0 d_\Gamma(M_0) + c_1\}.$$

On $\mathcal{E}_0$, fully evaluate the reserve. On its complement run the balanced reserve and delayed-activation policy of Appendix C.2, using only the chosen generating columns on these J_0 core types. Their envelope is exactly $\mathcal{F}_0^{\mathrm{net}}$ by (43). All later full and partial executions are charged as before.

The distance function is one-Lipschitz, so

$$d_\Gamma(M_0) \le d_0 + |M_0 - \mu_N|, \qquad \mathcal{E}_0 \subseteq \{|M_0 - \mu_N| \ge (m_0 - c_1)/c_0 - d_0\}.$$

$M_0 - \mu_N$ is the sum of centered exterior contributions. Their total variance is at most σ_N^2 and their ranges are at most $x_{\max}$. For independent centered Z_i with $|Z_i| \le b$ and $\sum_i \mathbb{E}Z_i^2 \le v$, Bernstein's inequality states

$$\Pr\left(\left|\sum_i Z_i\right| \ge t\right) \le 2\exp\left(-\frac{t^2}{2v + 2bt/3}\right), \qquad t > 0.$$

Here $Z_i = X_i - \mathbb{E}X_i$ ranges over the completed exterior, $b = x_{\max}$ and $v \le \sigma_N^2$. Hence

$$\Pr(\mathcal{E}_0) \le \Pr(|M_0 - \mu_N| \ge t) \le \min\{1, 2e^{-t^2/(2\sigma_N^2 + 2x_{\max}t/3)}\} = \delta_1(t).$$

For $t \le 0$ use $\delta_1(t) = 1$. Also

$$0 \le F_0(K) \le \Lambda_0(d_0 + |K - \mu_N|), \qquad \mathbb{E}F_0(K)^2 \le 2\Lambda_0^2(d_0^2 + \sigma_N^2). \tag{105}$$

$$\begin{aligned}\mathcal{L}_N(m_0) = \mathbb{E}[F_0(K)\mathbf{1}_{\mathcal{E}_0}] &\overset{\text{Cauchy–Schwarz}}{\le} \sqrt{\mathbb{E}F_0(K)^2}\sqrt{\Pr(\mathcal{E}_0)} \\ &\le \sqrt{2}\Lambda_0\sqrt{\sigma_N^2 + d_0^2}\sqrt{\delta_1(t)}.\end{aligned}$$

Let $u = 2\log(2 + \tau_N)$, $a = \sqrt{2\sigma_N^2 u}$ and $b = 2x_{\max}u/3$. For $t_u = a + b$,

$$\begin{aligned}t_u^2 - 2u(\sigma_N^2 + x_{\max}t_u/3) &= (a+b)^2 - a^2 - b(a+b) = ab \ge 0, \\ t \ge t_u \quad &\Longrightarrow \quad \delta_1(t) \le 2e^{-u} = \frac{2}{(2+\tau_N)^2}, \\ \mathcal{L}_N(m_0) &\le 2\Lambda_0\frac{\sqrt{\sigma_N^2 + d_0^2}}{2 + \tau_N} \le 2\Lambda_0.\end{aligned}$$

Under (K), Hoeffding's inequality gives, for sufficiently large N,

$$\begin{aligned}\mathcal{E}_0 &\subseteq \{|M_0 - \mathbb{E}K| \ge m_0/(2c_0)\}, \\ \Pr(\mathcal{E}_0) &\le 2\exp\left[-\frac{m_0^2}{2c_0^2 N {x_{\max}}^2}\right].\end{aligned} \tag{106}$$

Indeed $M_0 - \mathbb{E}K$ is the sum of the centered background aggregates; Hoeffding applies to these independent variables, whatever their response graphs or full costs. Also, for any genuine boundary k,

$$0 \le \mathcal{F}_0^{\mathrm{net}}(K) \le L_{\mathrm{net}}(|K - \mathbb{E}K| + c_{\mathrm{gate}}\sqrt{N} + 1), \qquad \mathbb{E}[\mathcal{F}_0^{\mathrm{net}}(K)^2] = O(N). \tag{107}$$

For $L_{\mathrm{core}} \ge 2c_0 x_{\max}$, Cauchy–Schwarz gives

$$\mathbb{E}[\mathcal{F}_0^{\mathrm{net}}(K)\mathbf{1}_{\mathcal{E}_0}] \le O(\sqrt{N})\Pr(\mathcal{E}_0)^{1/2} = O(1). \tag{108}$$

Only omitted net value is lost on this branch; no bound on the cost of the background's already-required full routines enters this estimate.

Conditional tracking and total fees. On $\mathcal{E}_0^c$, let $\mathscr{F}$ be the sigma-field of the final balanced transcript. It leaves m untouched roots per core type and $c_0 d_\Gamma(M) < m \le c_0 d_\Gamma(M) + c_1$, $M = \mathbb{E}[K \mid \mathscr{F}]$.

All the controller constants are those of the core. Conditional on this transcript, $K - M$ is a sum of $J_0 m$ independent bounded core variables. Write d_{knot} for the distance from this M to the core envelope's cell endpoints and basis changes, as in that lemma. Apply Lemma C with entry event $\mathcal{E}_0^c$. Its conditional bound and the strengthened averaging conclusion give

$$\mathbb{E}\Big[\mathbf{1}_{\mathcal{E}_0^c}\sqrt{m}e^{-c_{\mathrm{tr}}d_{\mathrm{knot}}^2/m}\Big] \le A'(1+\mathfrak{p}_N\mathfrak{K}_N). \tag{109}$$

On $\mathcal{E}_0$, set $m = m_0$ and $\mathscr{F} = \mathscr{F}_0$ before completing the reserve. On $\mathcal{E}_0^c$, retain the final balanced transcript. On both branches set $V_{\mathrm{fresh}} = m\sum_j v_j(\mathfrak{h}_j^0)$. Let C_{pre} be the fee up to this transcript. The remaining policy is full completion on $\mathcal{E}_0$ and conditional tracking on $\mathcal{E}_0^c$. Before the partial-policy stage every executed routine is minimizing, so

$$\begin{aligned}
\mathbb{E}C_{\mathrm{pre}} + \mathbb{E}V_{\mathrm{fresh}} &= V_0,\\
\mathbb{E}C_{\mathsf{Track}_0} &= \mathbb{E}C_{\mathrm{pre}} + \mathbb{E}[\mathbf{1}_{\mathcal{E}_0}V_{\mathrm{fresh}}] + \mathbb{E}[\mathbf{1}_{\mathcal{E}_0^c}\mathbb{E}(C_{\mathrm{remaining}} \mid \mathscr{F})]\\
&\le V_0 - \mathbb{E}[\mathbf{1}_{\mathcal{E}_0^c}\mathbb{E}(F_0(K) \mid \mathscr{F})]\\
&\quad + A\mathbb{E}[\mathbf{1}_{\mathcal{E}_0^c}(1+\sqrt{m}e^{-c_{\mathrm{tr}}d_{\mathrm{knot}}^2/m})].
\end{aligned}$$

The event $\mathcal{E}_0$ belongs to $\mathscr{F}$. Applying the tower property to the middle term gives

$$\begin{aligned}
\mathbb{E}C_{\mathrm{core}} &\le V_0 - \mathbb{E}[\mathcal{F}_0^{\mathrm{net}}(K)\mathbf{1}_{\mathcal{E}_0^c}] + A\mathbb{E}[\mathbf{1}_{\mathcal{E}_0^c}(1+\sqrt{m}e^{-c_{\mathrm{tr}}d_{\mathrm{knot}}^2/m})]\\
&\le V_0 - \mathbb{E}\mathcal{F}_0^{\mathrm{net}}(K) + \mathcal{L}_N(m_0) + A''(1+\mathfrak{p}_N\mathfrak{K}_N).
\end{aligned} \tag{110}$$

This is (100). Lemma B ensures exact output on every branch. Under (K), finitely many library knots have depth $O(\sqrt{N})$ and (63) gives $\mathfrak{p}_N = O(N^{-1/2})$. Equation (108) then gives the bounded remainder. □

C.4 Proof of the cost comparison and supply examples

The same lower accounting and reserve controller give the cost theorem without exact batch-price equality or the global count condition (K). For $q \ge 1$ distinct integer thresholds set

$$H_\Gamma = k_q - k_1, \qquad w_\Gamma = \max_{I \text{ internal}}(b_I - a_I), \qquad \zeta_N = \min\{w_\Gamma, \mathfrak{p}_N H_\Gamma\}, \tag{111}$$

where an empty maximum is zero and $\mathfrak{p}_N = \sup_z \Pr(K = z)$. The library contains at most M_{lib} selected columns. Its execution parameters satisfy the bounds stated in Theorem 1.

Proof of Theorem 1. Write $w_I = b_I - a_I$. The internal cells satisfy

$$\sum_{I \text{ internal}}(1 + w_I) = k_q - k_1 = H_\Gamma.$$

Each cell contributes at most M_{lib} internal ray intersections. The genuine boundary endpoints have zero depth. Thus the set used by the tracker satisfies

$$\begin{aligned}
\mathfrak{K}_N &\le 2q + \sum_{I \text{ internal}} M_{\mathrm{lib}}(1 + w_I) = 2q + M_{\mathrm{lib}}H_\Gamma\\
&\le 2 + 2H_\Gamma + M_{\mathrm{lib}}H_\Gamma = 2 + (M_{\mathrm{lib}} + 2)H_\Gamma,
\end{aligned}$$

Lemma D and the width bound in Lemma C give, simultaneously,

$$\begin{aligned}
\mathbb{E}C_{\mathsf{Track}_0} - (V_0 - \mathbb{E}F_0(K)) &\le \mathcal{L}_N(m_0) + C\min\{1 + \sqrt{w_\Gamma + 1}, 1 + \mathfrak{p}_N\mathfrak{K}_N\}\\
&\le \mathcal{L}_N(m_0) + C'\min\{1 + w_\Gamma, 1 + \mathfrak{p}_N H_\Gamma\}\\
&= \mathcal{L}_N(m_0) + C'(1 + \zeta_N).
\end{aligned}$$

For the lower comparison, (66) implies

$$\begin{aligned}
\mathfrak{B}_N &\le 2x_{\max}\Lambda_0\min\{w_\Gamma, (2x_{\max}+1)\mathfrak{p}_N H_\Gamma\}\\
&\le 2x_{\max}(2x_{\max}+1)\Lambda_0\zeta_N.
\end{aligned}$$

Set $c_{\mathrm{low}} = 2x_{\max}(2x_{\max}+1)\Lambda_0$ and let c_{up} be the last upper-bound constant. Then

$$\begin{aligned} V_0 - \mathbb{E}F_0(K) - \mathfrak{D}_N - c_{\mathrm{low}}\zeta_N &\le V_0 - \mathbb{E}F_0(K) - \mathfrak{D}_N - \mathfrak{B}_N \\ &\overset{(66)}{\le} C_N^* \le \mathbb{E}C_{\mathsf{Track}_0} \\ &\le V_0 - \mathbb{E}F_0(K) + \mathcal{L}_N(m_0) + c_{\mathrm{up}}(1+\zeta_N). \end{aligned}$$

The constants use only the selected column exposures, inverse bases, paid fees and completion prices. The theorem bounds these quantities uniformly. The local law itself may therefore vary between batch sizes. The Bernstein calculation in Lemma D proves the sufficient supply on the intrinsic scale $\tau_N = 1 + \sigma_N + d_\Gamma(\mu_N)$. □

Three cases follow from these inequalities. With one threshold, $H_\Gamma = w_\Gamma = \zeta_N = 0$, regardless of its position or the variance. With bounded internal widths, ζ_N is bounded even when q grows. Alternatively the finite concentration bound (63) gives $\mathfrak{p}_N \le C_x/(1+\sigma_N)$ for bounded integer contributions. Thus a total threshold span $H_\Gamma = O(1+\sigma_N)$ suffices, without requiring every threshold to be near the mean. In each case $\mathfrak{D}_N + \mathcal{L}_N(m_0) = O(1)$ yields the bounded expansion.

For a sparse-exception example, take $N-1$ independent fair bits with unit result-query fee as the library type, and one independent Bernoulli(p_N) bit with fee two. Its action is known. At the two library vertices $(2,0),(0,2)$ its Bellman shortfall is

$$d_{\mathrm{exceptional}} = 2 - 2\min(p_N, 1-p_N) \le 2.$$

All fair components have zero shortfall. For $p_N = N^{-2}$, $\min(p_N, 1-p_N) \to 0$ while $\mathfrak{D}_N \le 2$. Theorem 1 therefore gives a bounded single-threshold error when enough fair copies are available.

One family separates price accuracy from physical supply. Let N be odd and $k = (N+1)/2$. All actions are known. Each independent fair bit has a single result query: m components have fee r_h, and $N-m$ have fee r_l, where $0 < r_l < r_h$ and $1 \le m \le k-1$. Write T for first majority certification and $R = N - T$ for the number of unread bits. Every next unread bit is conditionally fair, in every adaptive query order. Therefore R has the same distribution in all orders. At first certification its exposure is $R/2$ and its overshoot is zero. Equation (48) gives

$$\mathbb{E}R = 2D_N, \qquad \Pr(R \ge u) = 2\Pr\{\mathrm{Binomial}(N-u, 1/2) \ge k\}, \quad 0 \le u \le k-1.$$

The two tail events correspond to observing k ones or k zeros in the first $N-u$ queries; they are disjoint and symmetric. Positive query fees allow restriction to policies stopping at the first certificate. For each realization the saved fee is at most $r_l R + (r_h - r_l)\min(R, m)$. Reading cheap components first attains that bound. In particular,

$$\begin{aligned} \sup_\pi \mathbb{E}[\text{saved fee}] &= \mathbb{E}[r_l R + (r_h - r_l)\min(R,m)] \\ &= r_h \mathbb{E}R - (r_h - r_l)\mathbb{E}(R-m)_+ = 2r_h D_N - (r_h - r_l)\mathbb{E}(R-m)_+. \end{aligned}$$

Subtracting this saving from V_0 gives

$$\begin{aligned} C_N^* &= V_0 - 2r_h D_N + (r_h - r_l)\mathbb{E}(R-m)_+, \qquad V_0 = r_l N + (r_h - r_l)m, \\ \mathbb{E}(R-m)_+ &= 2\sum_{u=m+1}^{k-1} \Pr\{\mathrm{Binomial}(N-u, 1/2) \ge k\}. \end{aligned} \tag{112}$$

The high-fee type alone generates the whole region $\mathcal{D} = \{\lambda \ge 0 : \lambda_+ + \lambda_- \ge 2r_h\}$. All product, count, support-anchor and fixed-core conditions hold for every $m \ge 1$. Its cone prediction misses exactly the final term in (112).

For $1 \le u \le k-1$ and $m \ge 1$,

$$\begin{aligned} \Pr(R \ge u) &= 2\Pr\{\mathrm{Binomial}(N-u, 1/2) - (N-u)/2 \ge (u+1)/2\} \\ &\overset{\text{Hoeffding}}{\le} 2\exp\left[-\frac{2((u+1)/2)^2}{N-u}\right] \le 2e^{-u^2/(2N)}, \\ \mathbb{E}(R-m)_+ &= \sum_{u=m+1}^{k-1} \Pr(R \ge u) \le 2\int_m^\infty e^{-x^2/(2N)}\,dx \\ &\le \frac{2}{m}\int_m^\infty x e^{-x^2/(2N)}\,dx = \frac{2N}{m}e^{-m^2/(2N)}. \end{aligned} \tag{113}$$

For a lower bound let $N \geq 65537$, $m \leq N/32$, $L = \lfloor \sqrt{N} \rfloor$, $u = m + L$, $n' = N - u$, and $j_0 = \lceil n'/2 \rceil$. These choices give $n' \geq N/2$. For every $j \in \{k, \dots, k + L - 1\}$, put $h = j - j_0$. Then $0 \leq h \leq m/2 + 2\sqrt{N} \leq n'/8$. If $p_j = 2^{-n'} \binom{n'}{j}$, its central mass and adjacent ratios satisfy

$$p_{j_0} \geq \frac{1}{2\sqrt{n'}}, \qquad \frac{p_{j_0+h}}{p_{j_0}} = \prod_{i=1}^{h} \left(1 - \frac{2j_0 - n' + 2i - 1}{j_0 + i}\right) \geq \exp[-4h(h+1)/n'].$$

The central-mass bound is the central-binomial inequality. For $z_i = (2j_0 - n' + 2i - 1)/(j_0 + i)$,

$$0 \leq z_i \leq 4i/n' \leq \tfrac{1}{2}, \qquad \log \frac{p_{j_0+h}}{p_{j_0}} = \sum_{i=1}^{h} \log(1 - z_i) \geq -2 \sum_{i=1}^{h} z_i \geq -\frac{8}{n'} \frac{h(h+1)}{2}.$$

Since $h + 1 \leq m + 3\sqrt{N}$, summing these L masses gives

$$\begin{aligned} \mathbb{E}(R - m)_+ &\geq L \Pr(R \geq m + L) = 2L \Pr\{\text{Binomial}(n', 1/2) \geq k\} \\ &\geq 2L \sum_{j=k}^{k+L-1} p_j \geq \frac{L^2}{\sqrt{n'}} e^{-72(1+m/\sqrt{N})^2} \\ &\geq \frac{\sqrt{N}}{4} e^{-72(1+m/\sqrt{N})^2}. \end{aligned} \tag{114}$$

Indeed $L \geq \sqrt{N}/2$, $L/\sqrt{n'} \geq 1/2$, and $4h(h+1)/n' \leq 72(1 + m/\sqrt{N})^2$. Thus $m = O(\sqrt{N})$ leaves a $\Theta(\sqrt{N})$ error; more generally, $m = o(\sqrt{N \log N})$ leaves an unbounded error for fixed $r_h - r_l > 0$. Conversely $m \geq \sqrt{N \log N}$ makes (113) at most $2/\sqrt{\log N}$. This family requires supply of order $\sqrt{N \log N}$ for a bounded remainder.

Finally set $m = k - 1$, so $R \leq m$ and the last term of (112) is zero. Choosing the abundant low-fee type as the proposed core gives $S_{0,N} = 2r_l D_N$, whereas the actual saving is $2r_h D_N$. Price-region equality fails and the error is $2(r_h - r_l)D_N = \Theta(\sqrt{N})$ for fixed distinct fees, despite linear supply of both types.

D Learning on the original batch

The known-law construction uses local response probabilities to choose queries. Here those probabilities are shared across components but unknown. We first derive feedback from purchased paths and bound the graph learner's regret. The reserve argument converts that regret into the extra query fee in Theorem 2. A two-sensor family then gives the matching lower rate.

D.1 Executed feedback and exact certification

For q copies of each type define the physical allocation value

$$\mathcal{H}_q(c) = \max_{z \geq 0} \left\{ \sum_{j,\pi} G_{j\pi} z_{j\pi} : \sum_{j,\pi} B^s_{j\pi} z_{j\pi} \leq c_s \ (s = +, -), \ \sum_{\pi} z_{j\pi} \leq q \ (j \leq J) \right\}. \tag{115}$$

At a known-action stopping history, executing a partial policy observes $(C_\pi, l(\mathfrak{h}_\pi), u(\mathfrak{h}_\pi)) \in [0, C_{\text{loc}}] \times [0, x_{\max}]^2$. By conditional expectation,

$$B^+_\pi = \mathbb{E}X - \mathbb{E}l(\mathfrak{h}_\pi), \quad B^-_\pi = \mathbb{E}u(\mathfrak{h}_\pi) - \mathbb{E}X, \quad G_\pi = v - \mathbb{E}C_\pi. \tag{116}$$

For a price $\lambda = (\lambda_+, \lambda_-)$ chosen before the execution,

$$\mathbb{E}C_\pi + \lambda \cdot B_\pi = \mathbb{E}[C_\pi - \lambda_+ l(\mathfrak{h}_\pi) + \lambda_- u(\mathfrak{h}_\pi)] + (\lambda_+ - \lambda_-)\mathbb{E}X. \tag{117}$$

The last term is common to the policies of a type. Thus their comparison uses only the executed cost and support endpoints.

Lemma L1 (Finite on-policy reduction)**.** *Conditional on the sigma-field $\mathcal{F}_0$ of a stopped transcript, let q untouched roots of each of J types remain, with all other outcomes summing to the known integer O. Choose a cell $[a, b]$ meeting the reachable count range and a center $x \in [a, b]$. Let $M_0 = O + q\sum_j \mathbb{E}X_j$. For type j, suppose an executed-path learner has expected priced regret at most $\mathfrak{R}_j(q)$ against every fixed policy mixture for any predictable sequence $\lambda_t \in [0, L_\lambda]^2$. Suppose graph-chosen full continuations satisfy* (69) *and $L_\lambda \geq \chi_{\mathrm{rec}}$. There is an exact reserve procedure satisfying*

$$\begin{aligned}\mathbb{E}[C_{\mathrm{reserve+repair}} \mid \mathcal{F}_0] \leq q\sum_j v_j - \mathcal{H}_q(x-a, b-x) + L_\lambda|x - M_0| + 2\chi_{\mathrm{rec}}x_{\max} \\ + \sum_j \mathfrak{R}_j(q) + 2L_\lambda J x_{\max}\sqrt{q}.\end{aligned} \tag{118}$$

A missing outer side, or a side already true for every remaining completion, is omitted in (115)*. The terminal-floor condition is included by choosing $\chi_{\mathrm{rec}} = C_{\mathrm{loc}}/b_*$, where b_* is that floor.*

Proof. In round t, select one partial policy per type at price λ_t, and observe

$$f_t = \sum_j C_{tj}, \qquad g_t = \big((a-O)/q - \sum_j l_{tj},\ \sum_j u_{tj} - (b-O)/q\big). \tag{119}$$

The selected cell guarantees $\|g_t\|_2 \leq \sqrt{2}Jx_{\max}$ on active sides. Set $\lambda_{t+1} = \mathrm{proj}_{[0,L_\lambda]^2}(\lambda_t + \gamma_{\mathrm{dual}}g_t)$, $\gamma_{\mathrm{dual}} = L_\lambda/(Jx_{\max}\sqrt{q})$. Inactive coordinates are zero. For any fixed $\lambda \in [0, L_\lambda]^2$, Euclidean projection gives

$$\begin{aligned}\|\lambda_{t+1} - \lambda\|_2^2 &\leq \|\lambda_t + \gamma_{\mathrm{dual}}g_t - \lambda\|_2^2 \\ &= \|\lambda_t - \lambda\|_2^2 + 2\gamma_{\mathrm{dual}}\langle\lambda_t - \lambda, g_t\rangle + \gamma_{\mathrm{dual}}^2\|g_t\|_2^2, \\ \sum_{t=1}^q \langle\lambda - \lambda_t, g_t\rangle &\leq \frac{\|\lambda_1 - \lambda\|_2^2 - \|\lambda_{q+1} - \lambda\|_2^2}{2\gamma_{\mathrm{dual}}} + \frac{\gamma_{\mathrm{dual}}}{2}\sum_{t=1}^q \|g_t\|_2^2 \\ &\leq \frac{L_\lambda^2}{\gamma_{\mathrm{dual}}} + \gamma_{\mathrm{dual}}qJ^2x_{\max}^2 = 2L_\lambda J x_{\max}\sqrt{q}.\end{aligned}$$

Choose $\lambda_s = L_\lambda \mathbf{1}_{\{\sum_t g_t^s > 0\}}$. This choice is allowed because the preceding inequality holds pathwise for every λ:

$$\sum_t f_t + L_\lambda\|[\sum_t g_t]_+\|_1 \leq \sum_t (f_t + \lambda_t \cdot g_t) + 2L_\lambda J x_{\max}\sqrt{q}. \tag{120}$$

For an optimal net allocation z, the comparator chooses π on type j with probability $z_{j\pi}/q$ and uses a minimizing full routine otherwise. For this mixture, let $\bar{f}, \bar{g}$ denote the expected one-round fee and constraint vector. With $c = (x - a, b - x)$,

$$\begin{aligned}q\bar{f} &= q\sum_j v_j - \sum_{j,\pi} z_{j\pi}G_{j\pi} = q\sum_j v_j - \mathcal{H}_q(c), \\ q\bar{g}_+ &= a - M_0 + \sum_{j,\pi} z_{j\pi}B_{j\pi}^+ \leq x - M_0, \\ q\bar{g}_- &= M_0 - b + \sum_{j,\pi} z_{j\pi}B_{j\pi}^- \leq M_0 - x, \\ \sum_t \mathbb{E}[\lambda_t \cdot \bar{g}] &\leq L_\lambda|x - M_0|.\end{aligned}$$

The last line uses nonnegative prices and the opposite signs of the two upper bounds. The assumed priced regret gives

$$\begin{aligned}\mathbb{E}\sum_t (f_t + \lambda_t \cdot g_t) &\leq \sum_t \mathbb{E}[\bar{f} + \lambda_t \cdot \bar{g}] + \sum_j \mathfrak{R}_j(q) \\ &\leq q\sum_j v_j - \mathcal{H}_q + L_\lambda|x - M_0| + \sum_j \mathfrak{R}_j(q).\end{aligned} \tag{121}$$

The comparator is fixed given $\mathcal{F}_0$. Each new root has its original law. At the end, $L = O + \sum_{t,j} l_{tj}$ and $U = O + \sum_{t,j} u_{tj}$, so

$$D := \left\| \left[\sum_t g_t \right]_+ \right\|_1 = (a - L)_+ + (U - b)_+.$$

The recourse lemma, with these L, U, a, b and the specified continuations, states $\mathbb{E}[C_{\text{repair}} \mid \mathscr{F}_q] \le \chi_{\text{rec}}(D + 2x_{\max})$. Hence

$$\begin{aligned}
\mathbb{E}C_{\text{reserve+repair}} &\le \mathbb{E}\sum_t f_t + \chi_{\text{rec}}\mathbb{E}D + 2\chi_{\text{rec}}x_{\max} \\
&\le \mathbb{E}\left[\sum_t f_t + L_\lambda D\right] + 2\chi_{\text{rec}}x_{\max} \\
&\overset{(120)}{\le} \mathbb{E}\sum_t (f_t + \lambda_t \cdot g_t) + 2L_\lambda J x_{\max}\sqrt{q} + 2\chi_{\text{rec}}x_{\max} \\
&\overset{(121)}{\le} q\sum_j v_j - \mathcal{H}_q + L_\lambda|x - M_0| + \sum_j \mathfrak{R}_j(q) + 2L_\lambda J x_{\max}\sqrt{q} + 2\chi_{\text{rec}}x_{\max}.
\end{aligned}$$

All expectations here are conditional on $\mathcal{F}_0$. □

Lemma L2 supplies the required regret bound by learning transition rows on the local history graph.

D.2 Learning on the query-history graph

A type has a layered acyclic legal-history graph with S_{hist} nodes, A_{row} query state–action rows, at most R_{reply} successors per row, and at most T_{path} paid queries per root. Its support, admissible stopping histories and $l(\mathfrak{h}), u(\mathfrak{h})$ are given; only the row probabilities $p(\cdot \mid \mathfrak{h}, e)$ are unknown. A root is a fresh independent mechanism of this type. Observing one path supplies its actual row transitions, without revealing counterfactual paths.

At episode t, a predictable price $\lambda_t \in [0, L_\lambda]^2$ is supplied. A query costs w_e; an action-sufficient stop has known terminal charge

$$\phi_t(\mathfrak{h}) = L_\lambda x_{\max} - \lambda_{t,+} l(\mathfrak{h}) + \lambda_{t,-} u(\mathfrak{h}) \in [0, 2L_\lambda x_{\max}]. \tag{122}$$

Thus the path score lies in $[0, C_{\text{score}}]$, $C_{\text{score}} = C_{\text{loc}} + 2L_\lambda x_{\max}$. The possible stage and terminal scores are known throughout the support graph. Only transition rows are learned. Episodic bandit-cost models study known transitions (Zimin and Neu, 2013) and unknown transitions (Jin et al., 2020). Here the observed costs and known terminal scores let us analyze the row estimates directly.

Let $n_t(\mathfrak{h}, e)$ be the number of row observations before the episode, and $\widehat{p}_t$ their empirical law. Intersect the legal-successor simplex with $\|p - \widehat{p}_t\|_1 \le \epsilon^{\text{row}}_{n_t(\mathfrak{h},e)}$, using the radius in (124). An unobserved row has the whole legal simplex. Denote this set by $\mathcal{P}_t(\mathfrak{h}, e)$. Let $V^\pi_p(\mathfrak{h})$ denote the true expected remaining score of policy π for the current episode's terminal charge, and let $U^*_t(\mathfrak{h}) = \min_\pi V^\pi_p(\mathfrak{h})$. Write $\mathbb{E}_t$ for expectation conditional on all observations before episode t. Backward evaluation gives

$$\widetilde{U}_t(\mathfrak{h}) = \min\left\{ \phi_t(\mathfrak{h}) \text{ if stopping is allowed}, \quad \min_e \left[w_e + \min_{p \in \mathcal{P}_t(\mathfrak{h},e)} p^T \widetilde{U}_t(\mathfrak{h}^{e,\cdot}) \right] \right\}. \tag{123}$$

Execute its minimizing policy under actual responses. The optimistic rows are only a value relaxation; they need not come from one common response mechanism. The output check still uses the entire true logical support. The row minimizer transfers probability mass from larger successor values to smaller ones, after sorting. Its arithmetic cost is $O(R_{\text{reply}} \log R_{\text{reply}})$; a backwards pass takes $O(A_{\text{row}} R_{\text{reply}} \log R_{\text{reply}} + S_{\text{hist}})$ and memory $O(A_{\text{row}} R_{\text{reply}} + S_{\text{hist}})$. Rational outward rounding of confidence radii and value minimization can add any specified summable optimization tolerance; exact output validity is unchanged.

D.3 Expected regret of the graph learner

Use the history-graph learner and predictable terminal prices of Appendix D.2. A type has S_{hist} histories, $A_{\mathrm{row}} \geq 1$ query rows, at most R_{reply} successors per row, and horizon T_{path}; put $C_{\mathrm{score}} = C_{\mathrm{loc}} + 2L_\lambda x_{\max}$. The following analysis controls model underestimation in expectation using a horizon-aware radius. The principle is related to minimax bandit confidence design (Audibert and Bubeck, 2009). The row estimates below allow prices to depend on preceding observations.

For an episode budget $q \geq 1$, let $T = 2^{\lceil \log_2 q \rceil} \in [q, 2q)$, $a = 1 + \log A_{\mathrm{row}} + R_{\mathrm{reply}} \log 2$, and use the same optimistic backward recursion, with

$$\epsilon_0^{\mathrm{row}} = 2, \qquad \epsilon_n^{\mathrm{row}} = \min\left\{2, \sqrt{\frac{32[a + \log(T/n)]}{n}}\right\} \quad (1 \leq n \leq q). \tag{124}$$

These radii change neither the observed feedback nor the permissible stops. Every unknown row is initially the full legal simplex; empirical zero counts never remove a response from the correctness support.

Lemma L2 (History-graph realization)**.** *Let $\widehat{p}_{en}$ be the first n observations of query row e, and set*

$$Z = \max_e \max_{1 \leq n \leq q} (\|\widehat{p}_{en} - p_e\|_1 - \epsilon_n^{\mathrm{row}})_+.$$

Unvisited stack entries are proof variables, not observations. Choose the actual maximum query depth $1 \leq T_{\mathrm{path}} \leq A_{\mathrm{row}}$ and put

$$\begin{aligned} a_{\mathrm{vis}} &= a + \log(2A_{\mathrm{row}}/T_{\mathrm{path}}), \\ \kappa_{\mathrm{graph}} &= 6T_{\mathrm{path}} + \min\big\{80A_{\mathrm{row}}\sqrt{a}, \\ &\qquad 2\sqrt{A_{\mathrm{row}}T_{\mathrm{path}}}[2 + \sqrt{32}(2\sqrt{a_{\mathrm{vis}}} + \sqrt{2\pi})]\big\}. \end{aligned} \tag{125}$$

Then

$$\mathbb{E}Z \leq 3q^{-1/2}, \qquad \mathrm{Reg}_{\mathrm{graph}}(q) \leq C_{\mathrm{score}}\kappa_{\mathrm{graph}}\sqrt{q}. \tag{126}$$

The regret is against the current true optimum at each predictable price. The planner takes $O(A_{\mathrm{row}}R_{\mathrm{reply}} \log R_{\mathrm{reply}} + S_{\mathrm{hist}})$ arithmetic operations and $O(A_{\mathrm{row}}R_{\mathrm{reply}} + S_{\mathrm{hist}})$ memory per episode and does not enumerate deterministic policies.

Proof. Successive visits to a row have its fixed conditional response law. Realize these visits by iid row stacks; neither independence between different stacks nor observations of unused stack entries are required by the argument. For any subset A of a row's response alphabet, $S_n(A) = \sum_{i=1}^n [\mathbf{1}\{X_i \in A\} - p_e(A)]$ is a bounded-increment centered sum. For $L = T2^{-j}$ and integer $n \in (L/2, L]$,

$$(\|\widehat{p}_{en} - p_e\|_1 - \epsilon_n^{\mathrm{row}})_+ > x \implies \max_A S_n(A) > \frac{L}{4}\left[x + \sqrt{\frac{32(a + j\log 2)}{L}}\right].$$

If the untruncated radius exceeds 2, the left event is empty, so the implication is still valid. For $\lambda > 0$ the centered indicator increments satisfy $\mathbb{E}[e^{\lambda(S_{n+1} - S_n)} \mid \mathscr{F}_n] \leq e^{\lambda^2/8}$. The exponential supermartingale gives

$$\Pr\{\max_{n \leq L} S_n(A) > z\} \leq \inf_{\lambda > 0} e^{-\lambda z + L\lambda^2/8} = e^{-2z^2/L},$$

$$\frac{2}{L}\left\{\frac{L}{4}\left[x + \sqrt{\frac{32(a + j\log 2)}{L}}\right]\right\}^2 \geq \frac{Lx^2}{8} + 4(a + j\log 2).$$

Union over at most $A_{\mathrm{row}}2^{R_{\mathrm{reply}}}$ subsets and over dyadic j yields

$$\begin{aligned} \Pr(Z > x) &\leq A_{\mathrm{row}}2^{R_{\mathrm{reply}}} \sum_{j=0}^{\log_2 T} e^{-4(a + j\log 2)}e^{-T2^{-j}x^2/8} \\ &\leq \sum_{j \geq 0} 2^{-4j}e^{-T2^{-j}x^2/8}, \\ \mathbb{E}Z &\leq \sqrt{\frac{2\pi}{T}} \sum_{j \geq 0} 2^{-7j/2} \leq \frac{3}{\sqrt{q}}. \end{aligned} \tag{127}$$

Only the inequality $\|\widehat{p} - p\|_1 = 2\max_A[\widehat{p}(A) - p(A)]$ is used; all actual row counts can be selected adaptively.

At round t, let Z_t be the same excess using only that round's current empirical rows. It is measurable before the new root and bounded by Z. Put $d_e = \|p_e - \widehat{p}_{te}\|_1$. For $d_e > 0$, define

$$\alpha_e = \min\{1, \epsilon^{\mathrm{row}}_{n_t(e)}/d_e\}, \qquad p'_e = (1-\alpha_e)\widehat{p}_{te} + \alpha_e p_e.$$

For $d_e = 0$, set $p'_e = p_e$. In either case,

$$\|p'_e - \widehat{p}_{te}\|_1 = \min(d_e, \epsilon^{\mathrm{row}}_{n_t(e)}), \qquad \|p'_e - p_e\|_1 = (d_e - \epsilon^{\mathrm{row}}_{n_t(e)})_+ \le Z_t.$$

The simulated model may be a relaxation of the complete-mechanism law; its policies and support are still legal. For a fixed policy π, put $\Delta(h) = V^\pi_{p'}(h) - V^\pi_p(h)$. At a stop, $\Delta(h) = 0$. At a query $e = \pi(h)$,

$$\begin{aligned}\Delta(h) &= [w_e + (p'_e)^T V^\pi_{p'}(h^{e,\cdot})] - [w_e + p_e^T V^\pi_p(h^{e,\cdot})] \\ &= (p'_e - p_e)^T V^\pi_p(h^{e,\cdot}) + (p'_e)^T \Delta(h^{e,\cdot}), \\ \Delta(h^0) &= \mathbb{E}_{p',\pi} \sum_{e \text{ queried}} (p'_e - p_e)^T V^\pi_p(h^{e,\cdot}).\end{aligned}$$

At most T_{path} rows are queried and every successor value lies in $[0, C_{\mathrm{score}}]$, so the absolute difference is at most $T_{\mathrm{path}} C_{\mathrm{score}} Z_t$. If π^* is true-optimal, optimistic minimization also gives $\widetilde{U}_t(\mathfrak{h}^0) \le V^{\pi^*}_{p'}(\mathfrak{h}^0)$. Subtracting $U^*_t(\mathfrak{h}^0) = V^{\pi^*}_p(\mathfrak{h}^0)$ proves

$$\widetilde{U}_t(\mathfrak{h}^0) - U^*_t(\mathfrak{h}^0) \le T_{\mathrm{path}} C_{\mathrm{score}} Z_t. \tag{128}$$

For the greedy optimistic policy, the minimizing row is $\widetilde{p}_{te}$. Subtract its optimistic Bellman equality from the true one and iterate along the actual path:

$$V^{\pi_t}_p(\mathfrak{h}^0) - \widetilde{U}_t(\mathfrak{h}^0) = \mathbb{E}_t \sum_{e \text{ visited}} (p_e - \widetilde{p}_{te})^T \widetilde{U}_t(\mathfrak{h}^{e,\cdot}).$$

The projected and optimistic rows are in the same radius ball:

$$\begin{aligned}\|p_e - \widetilde{p}_{te}\|_1 &\le \|p_e - p'_e\|_1 + \|p'_e - \widehat{p}_{te}\|_1 + \|\widehat{p}_{te} - \widetilde{p}_{te}\|_1 \\ &\le Z_t + 2\epsilon^{\mathrm{row}}_{n_t(e)}, \\ V^{\pi_t}_p(h^0) - U^*_t(h^0) &= [V^{\pi_t}_p(h^0) - \widetilde{U}_t(h^0)] + [\widetilde{U}_t(h^0) - U^*_t(h^0)] \\ &\le C_{\mathrm{score}} \mathbb{E}_t \sum_{e \text{ visited}} (2\epsilon^{\mathrm{row}}_{n_t(e)} + Z_t) + T_{\mathrm{path}} C_{\mathrm{score}} Z_t.\end{aligned}$$

Therefore

$$V^{\pi_t}_p(\mathfrak{h}^0) - U^*_t(\mathfrak{h}^0) \le 2C_{\mathrm{score}} \mathbb{E}_t \sum_{e \text{ visited}} \epsilon^{\mathrm{row}}_{n_t(e)} + 2T_{\mathrm{path}} C_{\mathrm{score}} Z_t. \tag{129}$$

This is conditional on the real past; Z itself is not assumed predictable. Taking expectations permits the uniform-stack bound on $\mathbb{E}Z_t$.

Each row is visited at most once per episode. For $1 \le v \le q$, write $b = a + \log(T/v)$. Monotonicity and $x = vu$ give

$$\begin{aligned}\sum_{n=0}^{v-1} \epsilon^{\mathrm{row}}_n &\le 2 + \sqrt{32} \int_0^v x^{-1/2} \sqrt{a + \log(T/x)}\, dx \\ &= 2 + \sqrt{32v} \int_0^1 u^{-1/2} \sqrt{b + \log(1/u)}\, du \\ &\le 2 + \sqrt{32v} \left[\sqrt{b} \underbrace{\int_0^1 u^{-1/2}\, du}_{2} + \underbrace{\int_0^1 u^{-1/2} \sqrt{\log(1/u)}\, du}_{\sqrt{2\pi}} \right].\end{aligned}$$

Here $\sqrt{x+y} \le \sqrt{x} + \sqrt{y}$ was used. Thus

$$\sum_{n=0}^{v-1} \epsilon_n^{\mathrm{row}} \le 2 + \sqrt{32v}\,[2\sqrt{a + \log(T/v)} + \sqrt{2\pi}] \le 40\sqrt{aq}.$$

For the last inequality, $v[a + \log(T/v)] \le Ta$ since $a \ge 1$ and $v \le T$. Summing the coarse bound over rows gives the coefficient $80A_{\mathrm{row}}\sqrt{a} + 6T_{\mathrm{path}}$. A sharper sum uses the actual visit counts v_e. Let $V_{\mathrm{vis}} = \sum_e v_e \le T_{\mathrm{path}} q$ and let A_{vis} count nonzero v_e. Since $A_{\mathrm{vis}} \le \min(A_{\mathrm{row}}, V_{\mathrm{vis}}) \le \sqrt{A_{\mathrm{row}} T_{\mathrm{path}} q}$, Cauchy–Schwarz and the log-sum inequality yield

$$\sum_e \sqrt{v_e[a + \log(T/v_e)]} \le \sqrt{A_{\mathrm{row}} V_{\mathrm{vis}}[a + \log(A_{\mathrm{row}} T/V_{\mathrm{vis}})]} \le \sqrt{A_{\mathrm{row}} T_{\mathrm{path}} q[a + \log(2A_{\mathrm{row}}/T_{\mathrm{path}})]},$$
$$\sum_e \sqrt{v_e} \le \sqrt{A_{\mathrm{row}} T_{\mathrm{path}} q}.$$

For $p_e = v_e/V_{\mathrm{vis}}$, the entropy bound $-\sum_e p_e \log p_e \le \log A_{\mathrm{row}}$ gives

$$\left(\sum_e \sqrt{v_e[a + \log(T/v_e)]}\right)^2 \overset{\text{Cauchy–Schwarz}}{\le} A_{\mathrm{row}} \sum_e v_e[a + \log(T/v_e)] = A_{\mathrm{row}} V_{\mathrm{vis}}[a + \log(T/V_{\mathrm{vis}}) - \sum_e p_e \log p_e] \le A_{\mathrm{row}} V_{\mathrm{vis}}[a + \log(A_{\mathrm{row}} T/V_{\mathrm{vis}})].$$

The function $x[a + \log(A_{\mathrm{row}}T/x)]$ is nondecreasing for $0 < x \le A_{\mathrm{row}}T$, since its derivative is $a - 1 + \log(A_{\mathrm{row}}T/x) \ge 0$. This proves the second inequality using $V_{\mathrm{vis}} \le T_{\mathrm{path}} q$ and $T < 2q$. The case $V_{\mathrm{vis}} = 0$ contributes zero. Therefore the total radius cost is

$$\sum_e \sum_{n < v_e} \epsilon_n^{\mathrm{row}} \le \sqrt{A_{\mathrm{row}} T_{\mathrm{path}} q}[2 + \sqrt{32}(2\sqrt{a_{\mathrm{vis}}} + \sqrt{2\pi})].$$

Write $B = 2 + \sqrt{32}(2\sqrt{a_{\mathrm{vis}}} + \sqrt{2\pi})$. Summing (129) gives

$$\mathrm{Reg}_{\mathrm{graph}}(q) \le 2C_{\mathrm{score}} \mathbb{E} \sum_e \sum_{n < v_e} \epsilon_n^{\mathrm{row}} + 2T_{\mathrm{path}} C_{\mathrm{score}} \sum_{t=1}^{q} \mathbb{E} Z_t \le C_{\mathrm{score}} \sqrt{q} \left[\min\{80A_{\mathrm{row}}\sqrt{a},\, 2\sqrt{A_{\mathrm{row}} T_{\mathrm{path}}} B\} + 6T_{\mathrm{path}}\right] = C_{\mathrm{score}} \kappa_{\mathrm{graph}} \sqrt{q}.$$

□

Applying the graph bound to the reserve. In the adaptive reserve of Theorem 2, its realized size m is determined before the fresh reserved roots. Restart row counts there and choose the horizon $q = m$ in (124). The conditional bound above holds at every predictable dual-price sequence, so it bounds the regret term in (142) by

$$C_{\mathrm{score}} \sum_j \kappa_{\mathrm{graph},j} \sqrt{\bar{m}_N}. \tag{130}$$

When $A_{\mathrm{row},j} = 0$, the type has zero query-choice regret. All constants here depend on the explicit graph size; its expansion from a succinct address description may be exponential.

From graph regret to query cost. For q fresh roots per type, replace the menu learner in (119) by this graph planner and retain the same projected prices and endpoint correction. Put $\kappa_{\mathrm{graph},j}$ as in

(125). For $L_\lambda \geq \chi_{\rm rec}$ under (69),

$$\mathbb{E}C_{\rm graph+repair} \leq q\sum_j v_j - \mathcal{H}_q(x-a, b-x) + L_\lambda|x-M_0| + 2\chi_{\rm rec}x_{\max} + \left[C_{\rm score}\sum_j \kappa_{{\rm graph},j} + 2L_\lambda J x_{\max}\right]\sqrt{q}. \tag{131}$$

For any feasible comparator mixture $\bar\pi_j$,

$$U_t^*(\mathfrak{h}_j^0) \leq V_p^{\bar\pi_j}(\mathfrak{h}_j^0), \qquad \mathbb{E}\sum_{t=1}^{q}[V_p^{\pi_{tj}}(\mathfrak{h}_j^0) - V_p^{\bar\pi_j}(\mathfrak{h}_j^0)] \leq C_{\rm score}\kappa_{{\rm graph},j}\sqrt{q}.$$

Thus the input in (118) is $\mathfrak{R}_j(q) = C_{\rm score}\kappa_{{\rm graph},j}\sqrt{q}$. Substitution gives the last line of (131). The fee counts queries and actual recourse. Terminal scores serve only to choose policies.

D.4 Reserve selection and the learning upper bound

Fix J legal core history graphs and fees, each with total path fee at most $C_{\rm loc}$. Each core type has at least $2s_N$ copies. Given the core membership and shared-type labels, designate s_N pilot roots and s_N reserve roots per type before observing responses. Every other root may have its own growing finite graph and law. A full routine is supplied as an explicit finite decision table for every root. The product model satisfies (P). Core laws are shared within type and unknown, and may vary with N on the given positive support. At each action-known unresolved history, choose a support-valid continuation from the graph and fees, for example by minimum worst-case fee with fixed tie breaking. For known $\chi_{\rm rec}, \beta_0 > 0$, require the completion-price condition and define

$$\begin{aligned} &\text{(L)} \qquad \mathbb{E}[C_{\kappa_j(\mathfrak{h})} \mid \mathfrak{h}] \leq \chi_{\rm rec} b_j^s(\mathfrak{h}) \quad (j \text{ core},\ \mathfrak{h} \in \mathcal{S}_j,\ s = +,-),\\ &\Pi_\mu^{\rm cov} = \{\pi \text{ in core} : G_\pi > 0,\ \min_s B_\pi^s \geq \beta_0\},\\ &\mathcal{V}_{0,\mu}(c) := \max_{z\geq 0}\left\{\sum_{\pi\in\Pi_\mu^{\rm cov}} G_\pi z_\pi : \sum_{\pi\in\Pi_\mu^{\rm cov}} B_\pi^s z_\pi \leq c_s\right\}. \end{aligned} \tag{132}$$

This library defines the comparator. Other legal policies and background components enter through its shortfall. Omit absent sides and suppose the covered library is nonempty. Denote its count envelope by F_0 and its axial-price sum by $\Lambda_0 \leq 2C_{\rm loc}/\beta_0$. Write $\mathfrak{M}_N^{\rm lib}$ for all product laws with the stated completion-price condition and library; background laws are arbitrary. Its heterogeneous shortfall $\mathfrak{D}_N$ and boundary correction $\mathfrak{B}_N$ are (65)–(66) for this library. The algorithm is not supplied these quantities or the identities of the covered policies. For comparison, let $\mathfrak{M}_N^{\rm cov} \subseteq \mathfrak{M}_N^{\rm lib}$ be the specialization with a core terminal floor $b_0 > 0$, $\chi_{\rm rec} = C_{\rm loc}/b_0$, (K), exact equality $\mathcal{V}_{0,\mu} = \mathcal{V}_{{\rm batch},\mu}$, and either (SA) everywhere or the global terminal floor in Lemma A. The finite comparison uses none of those additional conditions. A terminal floor is sufficient for the completion-price condition.

Public routines and the statistical experiment. Fix the structural descriptor: the support graphs, targets, fees, type labels, core membership, thresholds, supply, and coverage constants. Let Σ_g be the finite set of deterministic support-valid full decision tables on the graph of component g. A table specifies queries and stops, not a numerical value or unpurchased response. The joint instance class is

$$\mathfrak{E}_N^{\rm lib} = \left\{(\mu, \boldsymbol{\sigma}) : \mu \in \mathfrak{M}_N^{\rm lib},\quad \sigma_g \in \Sigma_g,\quad \mathbb{E}_{\mu_g} C_{\sigma_g} = v_{g,\mu}(\mathfrak{h}_g^0) \text{ for every } g\right\}. \tag{133}$$

Define $\mathfrak{E}_N^{\rm cov}$ by restricting the law to $\mathfrak{M}_N^{\rm cov}$. The routine tuple $\boldsymbol{\sigma}$ is part of the instance and may depend on μ, but is specified before the batch mechanisms are drawn. The initial observation is this entire tuple. With private seed U independent of the mechanisms, the experiment has law and filtration

$$P_{\mu,\boldsymbol{\sigma}} = \delta_{\boldsymbol{\sigma}} \otimes P_U \otimes \bigotimes_g \mu_g, \qquad \mathscr{F}_t = \sigma(\boldsymbol{\sigma}, U, e_1, R_1, \ldots, e_t, R_t), \qquad e_{t+1} \text{ is } \mathscr{F}_t\text{-measurable}. \tag{134}$$

The learner may use every feature of the public decision tables, including information about μ implied by their identity. Query fees charge each executed edge. The known-law oracle $C_N^*(\mu)$ uses the same

support and paid queries; it already knows μ, so these tables do not improve its optimum. For a fixed tuple define the fiber

$$\mathfrak{M}_N^{\mathrm{lib}}(\boldsymbol{\sigma}) = \{\mu : (\mu, \boldsymbol{\sigma}) \in \mathfrak{E}_N^{\mathrm{lib}}\}.$$

On this fiber the public message is constant. A distribution-independent optimal routine, as in the code and sensor examples, makes the entire stated response family lie in one such fiber.

Routine identity can convey information. For two independent bits with unit query fees and target $X_1 \vee X_2$, let their success probabilities be $(4/5, 1/5)$ under one law and $(1/5, 4/5)$ under the other. Reading i first and stopping on success costs $2 - \Pr(X_i = 1)$ in expectation. The two orders thus have fees $(6/5, 9/5)$ or $(9/5, 6/5)$. The unique optimal first query reveals which law supplied the routine. This information belongs to the public message in (134).

The upper proof below fixes an arbitrary joint pair. At that pair the initial message is deterministic and the unqueried mechanisms still have their original product law. Completing a previously unqueried component by its supplied table has expected fee $v_{g,\mu}(\mathfrak{h}_g^0)$. These are the only properties of the table used by the proof. At a partially observed retained history, endpoint correction uses a legal bounded-fee continuation chosen from the support graph, as allowed by Lemma B; a root-optimal table is not assumed optimal at an arbitrary off-path history. The resulting constants are uniform over all pairs in (133).

Lemma L3 (Intrinsic-scale upper reduction)**.** *For a pair in $\mathfrak{E}_N^{\mathrm{lib}}$, suppose each core type has predictable-price regret $\mathfrak{R}_j(q) \le c_j\sqrt{q}$ on executed paths. Put $s = s_N \ge 1$, $d_0 = d_\Gamma(\mathbb{E}K)$ and $\tau_N = 1 + \sqrt{\mathrm{Var}\, K} + d_0$. The same exact pilot-and-reserve algorithm satisfies*

$$\mathbb{E}_\mu C_{\mathcal{A}(\boldsymbol{\sigma})} - C_N^*(\mu) \le \mathfrak{D}_N + \mathfrak{B}_N + C\left(\sqrt{\tau_N} + \frac{\tau_N}{\sqrt{s}}\right) + \widetilde{\mathcal{L}}_N(s). \tag{135}$$

Here $\widetilde{\mathcal{L}}_N(s)$ is the expected library value lost on the initial full-completion branch. With $c = 8/\beta_0$ and $C_0 = 1 + cJx_{\max}$ it is at most

$$\sqrt{2}\Lambda_0\sqrt{\sigma_N^2 + d_0^2}\sqrt{\delta_2((s - C_0)/c - d_0)},$$

using (101)*. The constant is uniform in the core laws and public routines; it depends on the graph-regret coefficients, $J, x_{\max}, C_{\mathrm{loc}}, \chi_{\mathrm{rec}}, \beta_0$. In the class $\mathfrak{E}_N^{\mathrm{cov}}$ with $s_N \ge L_{\mathrm{core}}\sqrt{N\log(2N)}$ for sufficiently large fixed L_{core}, one recovers*

$$\sup_{(\mu,\boldsymbol{\sigma})\in\mathfrak{E}_N^{\mathrm{cov}}} [\mathbb{E}_\mu C_{\mathcal{A}(\boldsymbol{\sigma})} - C_N^*(\mu)] \le C_{\mathrm{cov}} N^{1/4}. \tag{136}$$

The specialization also permits the global-floor lower bound instead of the anchored shortfall comparison. The finite estimate retains mean error, supply loss, and mismatch with the complete optimal query class.

Proof. Fix any $(\mu, \boldsymbol{\sigma}) \in \mathfrak{E}_N^{\mathrm{lib}}$ and suppress the fixed public argument in the following expectations. *The second moment of the estimated center.* Write $s = s_N$, $c = 8/\beta_0$, $C_0 = 1 + cJx_{\max}$ and $L_\lambda = \chi_{\mathrm{rec}}$. Complete the background and the designated s pilot roots of every core type. Let $\widehat{p}_j$ be the pilot mean, $p_j = \mathbb{E}X_j$, and $e = \sum_j |\widehat{p}_j - p_j|$. Put $e_j = |\widehat{p}_j - p_j|$. The s pilot outcomes of a type are independent and lie in $[0, x_{\max}]$. Set $\delta_s = 2J\exp[-s/(2c^2J^2x_{\max}^2)]$. Then

$$\begin{aligned} \mathbb{E}e^2 = \mathbb{E}\left(\sum_j e_j\right)^2 &\le J\sum_j \mathbb{E}e_j^2 = J\sum_j \mathrm{Var}(\widehat{p}_j) \le \frac{J^2x_{\max}^2}{4s}, \\ \Pr(e > 1/(2c)) &\le \sum_j \Pr(e_j > 1/(2cJ)) \\ &\overset{\text{Hoeffding}}{\le} 2J\exp[-2s(2cJ)^{-2}/x_{\max}^2] = \delta_s. \end{aligned} \tag{137}$$

The event here is used only to bound a moment; the algorithm receives neither e nor a shrinking confidence radius. With $m = s - t$ untouched roots per type, let O_t be the observed outside count and set

$$\widetilde{M}_t = O_t + m\sum_j \widehat{p}_j, \qquad M_t = O_t + m\sum_j p_j = \mathbb{E}[K \mid \mathscr{F}_t].$$

If initially $s \le cd_\Gamma(\widetilde{M}_0) + C_0$, complete the reserve. Otherwise complete balanced rounds until the first

$$m \le cd_\Gamma(\widetilde{M}_t) + C_0. \tag{138}$$

At the first crossing $T > 0$, use $|\widetilde{M}_T - \widetilde{M}_{T-1}| \le Jx_{\max}$:

$$\begin{aligned} m_T = m_{T-1} - 1 &> c\, d_\Gamma(\widetilde{M}_{T-1}) + C_0 - 1 \\ &\ge c\, d_\Gamma(\widetilde{M}_T) + C_0 - 1 - cJx_{\max} = c\, d_\Gamma(\widetilde{M}_T). \end{aligned}$$

Together with the stopping condition this gives

$$cd_\Gamma(\widetilde{M}_T) < m \le cd_\Gamma(\widetilde{M}_T) + C_0. \tag{139}$$

Choose $x = \widetilde{M}_T$ in a count cell, or a neighboring attainable cell endpoint in a unit gap. Then $|x - \widetilde{M}_T| \le 1$ and $d_\Gamma(x) \le d_\Gamma(\widetilde{M}_T)$. The true covered comparator at x uses at most $d_\Gamma(x)/\beta_0 < m/8$ roots. It therefore fits the actual reserve regardless of the pilot's estimation error. Apply Lemma L1 for m rounds, then exact endpoint recourse. The game learns from fresh core roots and uses no background probability.

On the initial-full branch set $m = 0$, $x = M_T = K$ after completion. On the other branches, $|\widetilde{M}_T - M_T| \le me$. On $\{e \le 1/(2c)\}$,

$$\begin{aligned} m &\le c\, d_\Gamma(\widetilde{M}_T) + C_0 \le c\, d_\Gamma(M_T) + cme + C_0 \\ &\le c\, d_\Gamma(M_T) + m/2 + C_0, \\ m &\le 2c\, d_\Gamma(M_T) + 2C_0 \le 2c|M_T - \mu_N| + 2cd_0 + 2C_0. \end{aligned}$$

Let $\mathcal{G} = \{e \le 1/(2c)\}$. Conditional Jensen and the deterministic bound $m \le s$ give

$$\begin{aligned} \mathbb{E}(M_T - \mu_N)^2 = \mathbb{E}\{\mathbb{E}[K - \mu_N \mid \mathscr{F}_T]\}^2 &\le \mathbb{E}(K - \mu_N)^2 = \sigma_N^2, \\ \mathbb{E}m^2 = \mathbb{E}[m^2 \mathbf{1}_{\mathcal{G}}] + \mathbb{E}[m^2 \mathbf{1}_{\mathcal{G}^c}] &\\ \le 12c^2 \mathbb{E}(M_T - \mu_N)^2 + 12c^2 d_0^2 + 12C_0^2 + s^2 \Pr(\mathcal{G}^c) &\\ \le 12c^2(\sigma_N^2 + d_0^2) + 12C_0^2 + s^2 \delta_s. & \end{aligned}$$

Define

$$\begin{aligned} &M_{2,N} = 12c^2(\sigma_N^2 + d_0^2) + 12C_0^2 + s^2\delta_s, \\ &\mathbb{E}m^2 \le M_{2,N}, \qquad \bar{m}_N := \sqrt{M_{2,N}}, \quad \mathbb{E}m \le \bar{m}_N, \\ &a_N := 1 + \frac{Jx_{\max}}{2}\sqrt{M_{2,N}/s}, \qquad \mathbb{E}|x - M_T| \le 1 + \sqrt{\mathbb{E}m^2\, \mathbb{E}e^2} \le a_N, \\ &\mathbb{E}|K - x| \le \tfrac{x_{\max}}{2}\sqrt{J\bar{m}_N} + a_N. \end{aligned} \tag{140}$$

This Cauchy–Schwarz step allows the reserve size to depend on the pilot. Untouched roots retain their original independent laws at the stopping transcript, which proves the last line. On the exceptional mean-estimation event the deterministic bound is $m \le s$, not $m \le N$; this accounts for the $s^2\delta_s$ term. Put $c' = (2c^2J^2x_{\max}^2)^{-1}$. The uniform moment constants can be chosen explicitly:

$$\begin{aligned} s^2\delta_s = 2Js^2e^{-c's} &\le 2J \sup_{x \ge 0} x^2 e^{-c'x} = \frac{8J}{\exp(2)(c')^2}, \\ M_{2,N} &\le 12c^2(\sigma_N^2 + d_0^2) + 12C_0^2 + \frac{8J}{\exp(2)(c')^2} \\ &\le C_M \tau_N^2, \qquad C_M = 12c^2 + 12C_0^2 + \frac{8J}{\exp(2)(c')^2}, \\ a_N &\le 1 + \frac{Jx_{\max}\sqrt{C_M}}{2}\frac{\tau_N}{\sqrt{s}}. \\ \mathbb{E}m \le \bar{m}_N &= O(\tau_N). \end{aligned} \tag{141}$$

Initial supply loss and total costs. The initial estimate equals the completed background sum plus twice each pilot outcome. Its mean is μ_N. Its variance is at most $2\sigma_N^2$, because each pilot type has an

equal number of independent reserve copies in the original count; its individual ranges are at most $2x_{\max}$. Bernstein's inequality gives, with $t_s = (s - C_0)/c - d_0$,

$$\delta_{\text{init}} := \Pr\{s \le cd_\Gamma(\widetilde{M}_0) + C_0\} \le \delta_2(t_s).$$

Local fee telescoping and the chosen completion prices imply

$$G_\pi \le \mathbb{E}v_j(\mathfrak{h}_\pi) \le \mathbb{E}C_{\kappa_j(\mathfrak{h}_\pi)} \le \chi_{\text{rec}} B_\pi^s \qquad (s = +, -).$$

Consequently $\Lambda := \Lambda_0 \le 2\chi_{\text{rec}}$ for the reference library. No price bound is imposed on background policies. Define

$$T_{2,N} = 2\Lambda^2(\sigma_N^2 + d_0^2).$$

On the initial full-completion event $\mathcal{E}$,

$$\begin{aligned}\widetilde{\mathcal{L}}_N(s) &= \mathbb{E}[F_0(K)\mathbf{1}_{\mathcal{E}}] \\ &\le \sqrt{\mathbb{E}F_0(K)^2 \Pr(\mathcal{E})} \le \sqrt{T_{2,N}\delta_{\text{init}}}.\end{aligned}$$

The baseline V_0 includes every background full routine.

Let $\mathfrak{r}_N$ be any valid lower correction to $V_0 - \mathbb{E}F_0(K)$. Equation (66) permits $\mathfrak{r}_N = \mathfrak{D}_N + \mathfrak{B}_N$ for arbitrary backgrounds. In the exact price-dominance subclass the global-floor alternative of Lemma A can also be used. Let $\mathcal{E}$ denote the initial full-completion event and let C_{pre} be the fee before the final m fresh copies per type. Set $m = 0$ after full completion on $\mathcal{E}$ as above, and put $R = \sum_j c_j + 2L_\lambda J x_{\max}$. The required substitutions are

$$\begin{aligned}&\mathbb{E}C_{\text{pre}} + \mathbb{E}\left[m\sum_j v_j\right] = V_0, \\ &\mathcal{H}_m(x - a, b - x) \ge F_0(x) \quad \text{on } \mathcal{E}^c, \\ &\mathbb{E}|K - M_T| \le \mathbb{E}\sqrt{\mathrm{Var}(K \mid \mathscr{F}_T)} \le \tfrac{x_{\max}\sqrt{J}}{2}\mathbb{E}\sqrt{m} \le \tfrac{x_{\max}\sqrt{J}}{2}M_{2,N}^{1/4}, \\ &\mathbb{E}C_{\mathcal{A}} \overset{(118)}{\le} V_0 - \mathbb{E}[\mathbf{1}_{\mathcal{E}^c}F_0(x)] + L_\lambda\mathbb{E}|x - M_T| + R\mathbb{E}\sqrt{m} + 2\chi_{\text{rec}}x_{\max}, \\ &\mathbb{E}C_{\mathcal{A}} - C_N^* \le \mathfrak{r}_N + \underbrace{\mathbb{E}[\mathbf{1}_{\mathcal{E}}F_0(K)]}_{\widetilde{\mathcal{L}}_N(s)} + \mathbb{E}[\mathbf{1}_{\mathcal{E}^c}(F_0(K) - F_0(x))] \\ &\qquad + L_\lambda\mathbb{E}|x - M_T| + R\mathbb{E}\sqrt{m} + 2\chi_{\text{rec}}x_{\max} \\ &\qquad \le \mathfrak{r}_N + \widetilde{\mathcal{L}}_N(s) + \Lambda\mathbb{E}|K - M_T| + (\Lambda + L_\lambda)\mathbb{E}|x - M_T| \\ &\qquad\quad + R\mathbb{E}\sqrt{m} + 2\chi_{\text{rec}}x_{\max}.\end{aligned}$$

Insert $\mathbb{E}|x - M_T| \le a_N$, $\mathbb{E}\sqrt{m} \le M_{2,N}^{1/4}$ and $\widetilde{\mathcal{L}}_N(s) \le \sqrt{T_{2,N}\delta_{\text{init}}}$:

$$\begin{aligned}\mathbb{E}C_{\mathcal{A}} - C_N^* \le &\left[\frac{\Lambda x_{\max}\sqrt{J}}{2} + \sum_j c_j + 2L_\lambda J x_{\max}\right] M_{2,N}^{1/4} \\ &+ (\Lambda + L_\lambda)a_N + 2\chi_{\text{rec}}x_{\max} + \mathfrak{r}_N + \sqrt{T_{2,N}\delta_{\text{init}}}.\end{aligned} \tag{142}$$

Write the three coefficients in (142) as A_1, A_2, A_3, in their displayed order. They depend only on the stated local bounds and the regret coefficients. Since $\tau_N \ge 1$,

$$\begin{aligned}&M_{2,N}^{1/4} \le C_1\sqrt{\tau_N}, \qquad a_N \le 1 + C_2\tau_N/\sqrt{s}, \\ &\mathbb{E}C_{\mathcal{A}} - C_N^* \le \mathfrak{D}_N + \mathfrak{B}_N + \widetilde{\mathcal{L}}_N(s) \\ &\qquad + A_1C_1\sqrt{\tau_N} + A_2(1 + C_2\tau_N/\sqrt{s}) + A_3 \\ &\qquad \le \mathfrak{D}_N + \mathfrak{B}_N + \widetilde{\mathcal{L}}_N(s) + C(\sqrt{\tau_N} + \tau_N/\sqrt{s}).\end{aligned}$$

Here the initial-loss term is its exact value before replacing it by the Bernstein upper bound. Under (K), exact price dominance, the stated support alternative, and $s \ge L_{\text{core}}\sqrt{N\log(2N)}$ with $L_{\text{core}} > 2cx_{\max}$, $\mathfrak{r}_N = O(1)$, the supply term is bounded eventually, $M_{2,N}^{1/4} = O(N^{1/4})$ and $a_N = O(N^{1/4})$.

This proves (136). All failed-estimation branches still use legal queries and support-exact recourse. For smaller batches, (142) retains the actual reserve and initial-supply loss. The sufficient coefficient $17x_{\max}/\beta_0$ is a convenient strict choice above $2cx_{\max} = 16x_{\max}/\beta_0$. The pilot appears twice in the initial estimated count, giving squared-range sum at most $2Nx_{\max}^2$ here, versus $Nx_{\max}^2$ in the known-law reserve. The two supply coefficients follow from these different squared-range sums. For the intrinsic supply certificate set $u = 2\log(2+\tau_N)$, $a = 2\sigma_N\sqrt{u}$ and $b = 4x_{\max}u/3$. Then

$$(a+b)^2 - u[4\sigma_N^2 + 4x_{\max}(a+b)/3] = (a+b)^2 - a^2 - b(a+b) = ab \geq 0.$$

Thus $\delta_2(a+b) \leq 2e^{-u}$, and

$$s \geq \max\{\tau_N,\ c[d_0 + 2\sqrt{2}\sigma_N\sqrt{\log(2+\tau_N)} + \tfrac{8}{3}x_{\max}\log(2+\tau_N)] + C_0\}$$

gives

$$\begin{aligned}\widetilde{\mathcal{L}}_N(s) &\leq \sqrt{2}\Lambda_0\sqrt{\sigma_N^2 + d_0^2}\,\sqrt{2e^{-2\log(2+\tau_N)}} \\ &= 2\Lambda_0\frac{\sqrt{\sigma_N^2 + d_0^2}}{2+\tau_N} \leq 2\Lambda_0, \\ \tau_N/\sqrt{s} &\leq \sqrt{\tau_N}.\end{aligned}$$

Substitution in (135) gives excess fee $\mathfrak{D}_N + \mathfrak{B}_N + O(\sqrt{\tau_N})$. The algorithm takes any feasible s as input and executes the observed reserve rule (138). □

Procedure 2: $\mathsf{Learn}(\boldsymbol{\sigma})$**.** Input: support graphs, labels, fees, $\chi_{\mathrm{rec}}, \beta_0$, public full routines, and $2s_N$ copies per core type. Set $a_L = 8/\beta_0$, $b_L = 1 + a_L J x_{\max}$, and $L_\lambda = \chi_{\mathrm{rec}}$.

1. Designate s_N pilot and s_N reserve components per type. Complete the background and pilots using $\boldsymbol{\sigma}$; estimate each core mean by its pilot average $\widehat{p}_j$.
2. Use the same balanced reserve rule with $\widetilde{M} = O + m\sum_j \widehat{p}_j$ and constants a_L, b_L. An initial supply failure uses exact full completion.
3. Choose the cell at $\widetilde{M}$ (a neighboring endpoint in a unit gap). Restart row counts on the m fresh components per type. At each predictable price λ_t, run the optimistic graph recursion (123); record only executed responses, fees and terminal support endpoints.
4. Update λ_t by projected ascent on the observed two endpoint deficits, with step $L_\lambda/(Jx_{\max}\sqrt{m})$ as in (119). After these m rounds, use the graph-chosen continuations satisfying (69) to repair the actual support interval exactly.

Proof of Theorem 2. For type j set $L_\lambda = \chi_{\mathrm{rec}}$ and $c_j = (C_{\mathrm{loc}} + 2\chi_{\mathrm{rec}}x_{\max})\kappa_{\mathrm{graph},j}$. The graph bound is $\mathfrak{R}_j(q) \leq c_j\sqrt{q}$ for every predictable price sequence. Freshness of the stopped reserve verifies the input of Lemma L1. Equations (135) and (66) then give the explicit substitution

$$\begin{aligned}\mathbb{E}_\mu C_{\mathsf{Learn}(\boldsymbol{\sigma})} - C_N^*(\mu) &\leq \mathfrak{D}_N + \mathfrak{B}_N + C\left(\sqrt{\tau_N} + \frac{\tau_N}{\sqrt{s}}\right) + \widetilde{\mathcal{L}}_N(s) \\ &\leq \mathfrak{D}_N + 2x_{\max}(2x_{\max}+1)\Lambda_0\zeta_N + C\left(\sqrt{\tau_N} + \frac{\tau_N}{\sqrt{s}}\right) + \widetilde{\mathcal{L}}_N(s).\end{aligned}$$

The second line is the theorem with $c_{\mathrm{low}} = 2x_{\max}(2x_{\max}+1)\Lambda_0$. The price condition gives $\Lambda_0 \leq 2\chi_{\mathrm{rec}}$. All constants are uniform over the declared joint instances. □

D.5 Local costs and information in the two-sensor family

An equal-baseline pair of response laws. For one component, let $Y \sim \mathrm{Bernoulli}(1/2)$, and conditionally independent hints $H_j = Y \oplus E_j$. Their noise probabilities are

$$\varepsilon_{\mathrm{h},1} = 1/4 - sd, \qquad \varepsilon_{\mathrm{h},2} = 1/4 + sd, \qquad s \in \{-1, 1\}, \quad 0 < d \leq 1/100.$$

Queries of Y, H_1, H_2 cost $1, c_h, c_h$, respectively, where $c_h = 1/5$. All eight response vectors have positive probabilities bounded below by $1/64$. The required local action is constant and known. Full-target evaluation must query Y. Querying Y directly is the public optimal full-evaluation routine

under both laws, with $V_0 = N$. The batch target is $\mathbf{1}\{K \geq k\}$, $K = \sum_g Y_g$, $k = (N+1)/2$, for odd N.

Write $u = 1/4 - d$, $v = 1/4 + d$ and $\rho = (1 - 2c_h)/u = (3/5)/u$. For the $+$ price the local stopping fee is $\rho \Pr(Y = 1 \mid \mathfrak{h})$. The four two-hint posteriors follow from Bayes' rule:

$$\begin{aligned} p_{00} &= \frac{uv}{uv + (1-u)(1-v)} = \frac{1/16 - d^2}{5/8 - 2d^2} \geq \frac{1/16 - 10^{-4}}{5/8} > \frac{1}{16}, \\ p_{11} &= 1 - p_{00}, \\ \{p_{01}, p_{10}\} &= \left\{ \frac{1}{2} \pm \frac{d}{3/8 + 2d^2} \right\} \subset [.47, .53], \\ \rho p_{01}, \rho p_{10} &\geq 2.4(.47) > 1. \end{aligned}$$

Here $p_{ab} = \Pr(Y = 1 \mid H_1 = a, H_2 = b)$. After both hints, the priced value is $\min\{\rho p_{ab}, 1\}$. After a first zero, compare stopping, reading Y, and buying the other hint:

$$\begin{aligned} Q_{\text{second}|0} &= c_h + \Pr(H_{\text{other}} = 0 \mid H_{\text{first}} = 0)\rho p_{00} + \Pr(H_{\text{other}} = 1 \mid H_{\text{first}} = 0) \\ &= c_h + [uv + (1-u)(1-v)] \frac{\rho uv}{uv + (1-u)(1-v)} + (u + v - 2uv) \\ &= \frac{29}{40} + \frac{3}{5} d + 2d^2 > \frac{13}{20} \geq \rho v \geq \rho u = \frac{3}{5}, \\ U(0) &= \min\{\rho \varepsilon_{\text{first}}, 1, Q_{\text{second}|0}\} = \rho \varepsilon_{\text{first}}. \end{aligned}$$

After a first one, both second-hint branches have priced value 1:

$$U(1) = \min\{\rho(1 - \varepsilon_{\text{first}}), 1, c_h + 1\} = 1, \qquad \rho(1 - \varepsilon_{\text{first}}) \geq 2.4(.74) > 1.$$

Each first hint is fair. Therefore

$$\begin{aligned} Q_{\text{better}} &= c_h + \tfrac{1}{2}\rho u + \tfrac{1}{2} = c_h + \tfrac{1}{2}(1 - 2c_h) + \tfrac{1}{2} = 1, \\ Q_{\text{worse}} &= c_h + \tfrac{1}{2}\rho v + \tfrac{1}{2} = Q_{\text{better}} + \tfrac{1}{2}\rho(v - u) = 1 + \rho d. \end{aligned}$$

Including direct evaluation and stopping at the root gives

$$Q_Y = 1, \qquad Q_{\text{better}} = 1, \qquad Q_{\text{worse}} = 1 + \rho d, \qquad Q_{\text{stop}} = \rho/2 = 1 + \gamma_d, \quad \gamma_d \geq 1/5. \tag{143}$$

Reflecting Y, H_1, H_2 gives the same root values for the $-$ price. For any $r < \rho$, the better-hint policy satisfies

$$c_h + \tfrac{1}{2} r u + \tfrac{1}{2} = 1 - \tfrac{1}{2}(\rho - r)u < 1 = Q_Y.$$

Together with (143), this identifies both net-retention indices as ρ. The displayed posteriors and their reflections give $b_0 \leq \Pr(Y = 1 \mid \mathfrak{h}) \leq 1 - b_0$, $b_0 = 1/16$, at every unresolved history.

For the oracle comparison, use the better-hint column and its reflection. Their parameters satisfy

$$\begin{aligned} B^+ = u/2 \geq .12, \qquad B^- = (1-u)/2 \geq .375, \qquad 2.4 \leq \rho \leq 2.5, \\ C_{\text{loc}} = 7/5, \qquad \max_{h,s} \frac{1}{b^s(h)} \leq 16, \qquad (B^+)^{-1} \leq 25/3. \end{aligned}$$

There is one threshold. Each tracker therefore uses a scalar basis. In Lemma D, take $m_0 = N$, $J_0 = 1$ and $\beta_{\text{gen}} \geq .12$. Then $c_0 \leq 200/3$ and $c_1 \leq 203/3$. The initial mean $M_0 = N/2$ is in the central unit gap, so $d_\Gamma(M_0) = 0$. For $N \geq 69$,

$$\mathcal{E}_0 = \{N \leq c_1\} = \varnothing, \qquad \mathcal{L}_N(N) = 0, \qquad \mathfrak{K}_N = 2.$$

Let A be the uniform tracking constant for the displayed parameter bounds. The finite lower and upper comparisons give

$$\begin{aligned} N - \rho D_N &\overset{(64)}{\leq} C^*_{N,s} \\ &\overset{(100)}{\leq} N - \rho D_N + A(1 + 2\mathfrak{p}_N) \leq N - \rho D_N + 3A, \qquad N \geq 69. \end{aligned}$$

For odd $N < 69$, use $0 \leq C^*_{N,s} - (N - \rho D_N) \leq \rho D_N \leq (5/4)\sqrt{69}$. Set $A_0 = \max\{3A, (5/4)\sqrt{69}\}$. Thus

$$C^*_{N,s} = N - \rho D_N + \zeta_{N,s}, \quad 0 \leq \zeta_{N,s} \leq A_0, \qquad D_N = \mathbb{E}[(K - k)_+ + (k - 1 - K)_+]. \tag{144}$$

All-query regret and information budgets. For an exact batch algorithm, let U count final unresolved outcomes. Let U_0 count those with no queried hint. Write Q_{h} for the number of paid hint queries. Let T count roots whose first hint precedes their result query. Let $N_{\mathrm{bad},s}$ count those first-before-Y hints that use the worse sensor. Queries after Y is known are permitted. Let o be the nonnegative terminal overshoot on the certified side, as in (48).

Use the binary leave-one prices from (64). For either price direction, the worse first hint has local Bellman excess ρd, and stopping an untouched root has terminal excess γ_d, by (143). All remaining local gaps are nonnegative. Let $\mathcal{G}_g$ be the sum of the local Bellman gaps and the terminal priced gap on component g. For a query e at history h, write its gap as $\delta_e(h) = w_e + \mathbb{E}_s[U_\lambda(h, X_e) \mid h] - U_\lambda(h) \geq 0$. At a local stop, put $\delta_{\mathrm{stop}}(h) = \lambda \cdot b(h) - U_\lambda(h) \geq 0$. With $U_\lambda(\varnothing) = 1$, telescoping each stopped local tree gives

$$\begin{aligned}
\mathbb{E}_s[C_{\mathcal{A},g} + \lambda_g \cdot b_g^{\mathrm{end}} - 1 \mid \omega_{-g}] &= \mathbb{E}_s\left[\sum_{t:\, g_t = g} \delta_{e_t}(h_t) + \delta_{\mathrm{stop}}(h_g^{\mathrm{end}}) \,\middle|\, \omega_{-g}\right] = \mathbb{E}_s[\mathcal{G}_g \mid \omega_{-g}], \\
\mathbb{E}_s C_{\mathcal{A}} - [N - \rho D_N] &= \mathbb{E}_s \sum_g \mathcal{G}_g + \rho D_N - \mathbb{E}_s \sum_g \lambda_g \cdot b_g^{\mathrm{end}} \\
&\overset{(48)}{=} \mathbb{E}_s \sum_g \mathcal{G}_g + \rho \mathbb{E}_s o \\
&\geq \rho d\, \mathbb{E}_s N_{\mathrm{bad},s} + \gamma_d \mathbb{E}_s U_0 + \rho \mathbb{E}_s o.
\end{aligned}$$

Since $C^*_{N,s} - [N - \rho D_N] \leq A_0$,

$$R_s + A_0 \geq \rho d\, \mathbb{E}_s N_{\mathrm{bad},s} + \gamma_d \mathbb{E}_s U_0 + \rho \mathbb{E}_s o, \qquad R_s = \mathbb{E}_s C_{\mathcal{A}} - C^*_{N,s} \geq 0. \tag{145}$$

The price can depend on the other realized components; the displayed local excesses hold for both directions and hence survive that projection. Put $D(K) = (K-k)_+ + (k-1-K)_+$ and $m_T = \mathbb{E}_s[K \mid \mathsf{R}_T]$. Let b_g^{cert} be the exposure on the certified side. Then

$$\begin{aligned}
b_0 U \leq \sum_g b_g^{\mathrm{cert}} &= \mathbb{E}_s[D(K) \mid \mathsf{R}_T] - o \leq \mathbb{E}_s[D(K) \mid \mathsf{R}_T], \\
b_0 \mathbb{E}_s U \leq D_N &= \mathbb{E}_s(|K - N/2| - 1/2) \leq \sqrt{\mathrm{Var}_s\, K} = \sqrt{N}/2, \\
\mathbb{E}_s|K - m_T| \leq \sqrt{\mathbb{E}_s(K-m_T)^2} &= \sqrt{\mathbb{E}_s \sum_g \mathrm{Var}_s(Y_g \mid \mathsf{R}_T)} \leq \tfrac{1}{2}\sqrt{\mathbb{E}_s U} \leq \tfrac{1}{2}\sqrt{D_N/b_0}.
\end{aligned}$$

Every result is queried at most once. Counting its unit fee and all hint fees gives

$$\begin{aligned}
C_{\mathcal{A}} &= N - U + c_h Q_{\mathrm{h}}, \\
c_h \mathbb{E}_s Q_{\mathrm{h}} &= R_s + \zeta_{N,s} - \rho D_N + \mathbb{E}_s U \leq R_s + A_0 + D_N/b_0.
\end{aligned} \tag{146}$$

Thus low regret limits every paid hint, including hints bought after Y.

A whole-response relative-entropy potential gives a tighter all-query information budget. For Bernoulli relative entropy, $\partial_q \,\mathrm{kl}(p,q) = (q-p)/[q(1-q)]$. Hence

$$\begin{aligned}
I_{\mathrm{full}}(d) &= \mathrm{kl}(u,v) + \mathrm{kl}(v,u) \\
&= \int_u^v \frac{q-u}{q(1-q)}\, dq + \int_u^v \frac{v-q}{q(1-q)}\, dq \\
&\leq \frac{1}{u(1-u)} \int_u^v (v-u)\, dq = \frac{4d^2}{u(1-u)} \leq \frac{4d^2}{.24(.76)} < 22d^2.
\end{aligned}$$

A root is activated when its first hint is purchased. Let Q_{first} count these activations. For an activated root define

$$\Psi_g(\mathfrak{h}_g) = \mathrm{KL}(\mu_+(\cdot \mid \mathfrak{h}_g) \| \mu_-(\cdot \mid \mathfrak{h}_g));$$

set its accounting potential to zero before activation. The unactivated history is either empty or contains only Y, whose distribution is the same under both laws. In either case the conditional KL of the full remaining response vector is exactly $I_{\mathrm{full}}(d)$, because the two sensor noises are conditionally independent given Y.

Let D be the complete query transcript, including private randomization. Let $\tau_{\mathrm{end}} \le 3N$ count its queries. Let i_t be the conditional KL of the next response, and let J_t indicate a first-hint activation. Write Ψ_t for the sum of activated-root potentials. For a local query the finite chain rule is

$$\mathrm{KL}(\mu_+(\omega_g \mid h)\|\mu_-(\omega_g \mid h)) = \sum_x p_+(x \mid h) \log \frac{p_+(x \mid h)}{p_-(x \mid h)} + \sum_x p_+(x \mid h)\, \mathrm{KL}(\mu_+(\omega_g \mid h, x)\|\mu_-(\omega_g \mid h, x)).$$

Before activation the full remaining KL equals $I_{\mathrm{full}}(d)$. An outcome-only response has zero KL. The three cases give

$$i_t + \mathbb{E}_+[\Psi_{t+1} - \Psi_t \mid \mathscr{F}_t] = I_{\mathrm{full}}(d) J_t.$$

The policy's query probabilities cancel in the transcript likelihood ratio. Since the number of distinct queries is bounded,

$$\begin{aligned} \mathrm{KL}(\mathbb{P}_+^D \| \mathbb{P}_-^D) &= \mathbb{E}_+ \sum_{t<\tau_{\mathrm{end}}} i_t \\ &= I_{\mathrm{full}}(d) \mathbb{E}_+ \sum_{t<\tau_{\mathrm{end}}} J_t - \mathbb{E}_+ \sum_{t<\tau_{\mathrm{end}}} (\Psi_{t+1} - \Psi_t) \\ &= I_{\mathrm{full}}(d) \mathbb{E}_+ Q_{\mathrm{first}} - \mathbb{E}_+ \Psi_{\tau_{\mathrm{end}}}. \end{aligned}$$

Using $\Psi_{\tau_{\mathrm{end}}} \ge 0$, $Q_{\mathrm{first}} \le Q_{\mathrm{h}}$, and Pinsker's inequality $\mathrm{TV}(P, Q)^2 \le \mathrm{KL}(P\|Q)/2$ gives

$$\begin{aligned} \mathrm{KL}(\mathbb{P}_+^D \| \mathbb{P}_-^D) = I_{\mathrm{full}}(d) \mathbb{E}_+ Q_{\mathrm{first}} - \mathbb{E}_+ \sum_g \Psi_g(\mathfrak{h}_g^{\mathrm{end}}) \le 22 d^2 \mathbb{E}_+ Q_{\mathrm{h}}, \\ \mathrm{TV}(\mathbb{P}_+^D, \mathbb{P}_-^D) \le \sqrt{11 d^2 \mathbb{E}_+ Q_{\mathrm{h}}}. \end{aligned} \tag{147}$$

The transcript includes algorithmic randomization, all resolved-root calibration and every interleaved query.

D.6 The fourth-root lower bound and matching upper rate

Corollary 1 (A shared two-sensor family)**.** *Use the family in Appendix D.5, with common public full routine $\boldsymbol{\sigma}^0$ that reads each result directly. There are constants $0 < c < C < \infty$ and N_0 such that, for every odd $N \ge N_0$,*

$$cN^{1/4} \le \inf_{\mathcal{A}} \sup_{|\theta| \le 1/100} \left[\mathbb{E}_\theta C_{\mathcal{A}(\boldsymbol{\sigma}^0)} - C_N^*(\theta)\right] \le CN^{1/4}. \tag{148}$$

All original query orders are permitted. Hints bought after a result query are included in the fee and information bounds.

Proof. Write $R_s = \mathbb{E}_s C_{\mathcal{A}} - C_{N,s}^*$ and $R = \max(R_+, R_-)$. Use the local constants $b_0 = 1/16$, $c_h = 1/5$, $\rho \ge 12/5$ and $\gamma_d \ge 1/5$. The oracle remainder is bounded by the uniform A_0 in (144).

A count margin of order $\sqrt{N}$. Let $W = K - N/2$ and $D(K) = |W| - 1/2$. For odd N, $D(K) \ge 0$ and

$$\mathbb{E}W^2 = N/4, \qquad \mathbb{E}W^4 = (3N^2 - 2N)/16 \le 3N^2/16.$$

For any nonnegative Z and $a < \mathbb{E}Z$, Cauchy–Schwarz gives

$$\begin{aligned} \mathbb{E}Z &= \mathbb{E}[Z\mathbf{1}_{\{Z<a\}}] + \mathbb{E}[Z\mathbf{1}_{\{Z\ge a\}}] \\ &\le a + \sqrt{\mathbb{E}Z^2 \Pr(Z \ge a)}, \qquad \Pr(Z \ge a) \ge \frac{(\mathbb{E}Z - a)^2}{\mathbb{E}Z^2}. \end{aligned}$$

Use $Z = W^2$ and $a = N/8$. For $N \ge 1152$,

$$\begin{aligned} \Pr(|W| \ge \sqrt{N/8}) &\ge \frac{(N/8)^2}{3N^2/16} = \frac{1}{12}, \\ D_N = \mathbb{E}D(K) &\ge \frac{\sqrt{N}}{24\sqrt{2}} - \frac{1}{2} \ge a_D \sqrt{N}, \qquad a_D := \frac{1}{48\sqrt{2}}, \\ D_N &\le \sqrt{\mathbb{E}W^2} = \frac{\sqrt{N}}{2}, \qquad \mathbb{E}D(K)^2 \le \mathbb{E}W^2 = N/4. \end{aligned}$$

Many paid opportunities are necessary. Let U count terminal unread results, U_0 those without any hint, and T the number of first hints bought before their results. Let $\mathscr{F}_{\mathrm{end}}$ contain the complete terminal record. Conditional exposure balance and the posterior floor give, pathwise,

$$b_0 U \le \mathbb{E}_s[D(K) \mid \mathscr{F}_{\mathrm{end}}], \qquad \mathbb{E}_s U^2 \le b_0^{-2}\mathbb{E}_s\{\mathbb{E}_s[D(K) \mid \mathscr{F}_{\mathrm{end}}]\}^2 \overset{\text{Jensen}}{\le} 64N.$$

Put $V = U - U_0$. Every hinted unread result contributes a first-before-result hint, so $0 \le V \le T$. The exact fee identity and regret budget imply

$$\begin{aligned}
&\mathbb{E}_s U = c_h \mathbb{E}_s Q_{\mathrm{h}} + \rho D_N - \zeta_{N,s} - R_s \ge \rho D_N - A_0 - R_s,\\
&\mathbb{E}_s U_0 \le (R_s + A_0)/\gamma_d \le 5(R + A_0),\\
&\mathbb{E}_s V \ge (12/5) a_D \sqrt{N} - 6(R + A_0), \qquad \mathbb{E}_s V^2 \le 64N.
\end{aligned}$$

If $R > N^{1/4}$, the required lower bound already holds. Otherwise choose N_0 large enough that $N^{1/4} + A_0 \le \min\{a_D\sqrt{N}/5, \sqrt{N}\}$ for $N \ge N_0$. Set $a_V = 6a_D/5$, $M = 128/a_V$, and $m = \lceil M\sqrt{N} \rceil$. Then

$$\begin{aligned}
\mathbb{E}_s \min(T, m) &\ge \mathbb{E}_s \min(V, m) = \mathbb{E}_s V - \mathbb{E}_s (V - m)_+\\
&\ge a_V \sqrt{N} - \frac{\mathbb{E}_s V^2}{m} \ge (a_V - 64/M)\sqrt{N} = \frac{a_V}{2}\sqrt{N}.
\end{aligned}$$

Here $(v - m)_+ \le v^2/m$ for $v \ge 0$. This controls the stopped opportunities without fixing their horizon.

The transcript cannot identify the better sensor cheaply. The all-query fee budget gives

$$\mathbb{E}_+ Q_{\mathrm{h}} \le 5(R + A_0 + D_N/b_0) \le 5(\sqrt{N} + 8\sqrt{N}) = 45\sqrt{N}.$$

Choose

$$a = \min\left\{\frac{1}{100}, \frac{a_V}{4(M+1)\sqrt{495}}\right\}, \qquad d = aN^{-1/4}.$$

Let P_+, P_- denote the complete transcript laws. Equation (147) gives

$$\begin{aligned}
&\mathrm{KL}(P_+ \| P_-) \le 22 d^2 \mathbb{E}_+ Q_{\mathrm{h}} \le 990 a^2,\\
&\mathrm{TV}(P_+, P_-) \le \sqrt{\mathrm{KL}(P_+ \| P_-)/2} \le \sqrt{495}\, a \le \frac{a_V}{4(M+1)}.
\end{aligned}$$

On a transcript, let B_s^m count wrong-sensor choices among the first $\min(T, m)$ opportunities under sign s. The algorithm has the same public input under both signs, and

$$0 \le B_s^m \le m, \qquad B_+^m + B_-^m = \min(T, m).$$

For $0 \le f \le m$, integration of its level sets gives

$$|\mathbb{E}_+ f - \mathbb{E}_- f| = \left| \int_0^m [P_+(f > t) - P_-(f > t)]\, dt \right| \le m\mathrm{TV}(P_+, P_-).$$

Consequently

$$\begin{aligned}
\mathbb{E}_+ N_{\mathrm{bad},+} + \mathbb{E}_- N_{\mathrm{bad},-} &\ge \mathbb{E}_+ B_+^m + \mathbb{E}_- B_-^m\\
&\ge \mathbb{E}_+[B_+^m + B_-^m] - m\mathrm{TV}(P_+, P_-)\\
&\ge \frac{a_V}{2}\sqrt{N} - (M+1)\sqrt{N}\frac{a_V}{4(M+1)} = \frac{a_V}{4}\sqrt{N},\\
2R + 2A_0 &\ge R_+ + R_- + 2A_0\\
&\overset{(145)}{\ge} \rho d[\mathbb{E}_+ N_{\mathrm{bad},+} + \mathbb{E}_- N_{\mathrm{bad},-}]\\
&\ge \frac{12}{5} a N^{-1/4} \frac{a_V}{4}\sqrt{N} = \frac{3}{5} a a_V N^{1/4}.
\end{aligned}$$

Increase N_0 so that $A_0 \le (3aa_V/20)N^{1/4}$. Then $R \ge (3aa_V/20)N^{1/4}$. Since $\theta = \pm d$ belongs to $[-1/100, 1/100]$, infimizing over all exact algorithms preserves this lower bound.

The matching upper rate. The same family has one core type, common full routine, fixed graph, posterior floor b_0, zero library defect and one threshold. Thus $\zeta_N = 0$ and $\tau_N = 1 + \sqrt{N}/2$. Choose $s = \lceil L\sqrt{N\log(2N)}\rceil$ with fixed sufficiently large L; $2s \le N$ for all sufficiently large N. The supply and learning bounds give

$$\widetilde{\mathcal{L}}_N(s) \le 2\Lambda_0, \qquad \sqrt{\tau_N} \le \sqrt{3/2}\, N^{1/4},$$
$$\frac{\tau_N}{\sqrt{s}} \le \frac{3N^{1/4}}{2\sqrt{L}\,[\log(2N)]^{1/4}},$$
$$\mathbb{E}_\theta C_{\mathsf{Learn}(\boldsymbol{\sigma}^0)} - C_N^*(\theta) \le c_{\mathrm{learn}} \left(\sqrt{3/2} + \frac{3}{2\sqrt{L}(\log 2)^{1/4}}\right) N^{1/4} + 2\Lambda_0 \le CN^{1/4}.$$

All constants are uniform for $|\theta| \le 1/100$. □

E FINITE QUERY COSTS AND RECORDED EVIDENCE

The finite calculations evaluate the two sources of cost in the main results: acquiring useful evidence and learning how to use it. We give the exact dynamic program first, followed by the learning and fixed-record protocols used in Figures 1 and 2.

E.1 EXACT FINITE COSTS IN THE ONE-HINT INTERFACE

Each component has a fair result bit Y and a hint $H = Y \oplus E$, with $E \sim \mathrm{Bernoulli}(\varepsilon_h)$. The result and hint fees are r and c_h. All local actions are known. The batch accepts when $K \ge k$.

Let u be the number of unqueried components. Let n_0, n_1 count unresolved components whose hints are zero and one. Let q be the additional number of passes needed for acceptance. Define $V(u, n_0, n_1; q)$ as the minimum expected remaining query fee. Exact stopping and the posterior laws give

$$V(u, n_0, n_1; q) = 0 \quad \text{if } q \le 0 \text{ or } q > u + n_0 + n_1,$$
$$\Pr(Y = 1) = \tfrac{1}{2}, \qquad \Pr(Y = 1 \mid H = 0) = \varepsilon_h, \qquad \Pr(Y = 1 \mid H = 1) = 1 - \varepsilon_h.$$

At every other state, a legal first query has fee w_e and reply R_e. Conditioning on that reply gives

$$\mathbb{E}C_e = w_e + \sum_x \Pr(R_e = x)\, V(\text{state after } R_e = x).$$

Thus the four available query choices have values

$$\begin{aligned}
Q_{\mathrm{direct}} &= r + \tfrac{1}{2}V(u-1, n_0, n_1; q-1) + \tfrac{1}{2}V(u-1, n_0, n_1; q),\\
Q_0 &= r + \varepsilon_h V(u, n_0 - 1, n_1; q-1) + (1-\varepsilon_h)V(u, n_0 - 1, n_1; q),\\
Q_1 &= r + (1-\varepsilon_h)V(u, n_0, n_1 - 1; q-1) + \varepsilon_h V(u, n_0, n_1 - 1; q),\\
Q_{\mathrm{hint}} &= c_h + \tfrac{1}{2}V(u-1, n_0 + 1, n_1; q) + \tfrac{1}{2}V(u-1, n_0, n_1 + 1; q),\\
V(u, n_0, n_1; q) &= \min\{Q_{\mathrm{direct}}, Q_0, Q_1, Q_{\mathrm{hint}}\}.
\end{aligned} \tag{149}$$

Omit a choice when its queried population is zero. A hint bought after its result is known has positive fee and changes no remaining conditional law. Removing it lowers the cost. The four choices therefore exhaust useful first queries. Result queries reduce $u + n_0 + n_1$; hint queries reduce u at fixed total. The recurrence is acyclic and

$$C_N^* = V(N, 0, 0; k), \qquad V_0 = rN.$$

For the plotted majority batches, N is odd and $k = (N+1)/2$. Put

$$D_N = 2^{-N} \sum_{j=0}^{N} \binom{N}{j} [(j-k)_+ + (k-1-j)_+].$$

Let R_T be the first result-only certificate and U_{result} its unread count. Each unread bit has exposure $1/2$. The first certificate has zero overshoot, so

$$\begin{aligned} \tfrac{1}{2}U_{\mathrm{result}} &\overset{(48)}{=} \mathbb{E}[(K-k)_+ + (k-1-K)_+ \mid \mathsf{R}_T], \\ \tfrac{1}{2}\mathbb{E}U_{\mathrm{result}} &= D_N, \\ V_0 - C_{\mathrm{result}} &= rN - r(N - \mathbb{E}U_{\mathrm{result}}) = 2rD_N. \end{aligned}$$

The all-query net-price lower bound yields

$$S_N = \widehat{\rho} D_N, \qquad V_0 - C_N^* \le S_N, \qquad \widehat{\rho} = \max\{2r, (r-2c_h)/\varepsilon_h\}.$$

At $(r, c_h, \varepsilon_h) = (10, 1, 0.1)$, Figure 1a uses all 401 archived odd-N values through $N = 801$. At that size,

$$\begin{aligned} V_0 &= 8010, & V_0 - C_N^* &= 744.240, \\ V_0 - C_{\mathrm{result}} &= 215.887, & S_N &= 863.549. \end{aligned}$$

The figure compares exact query fees and a net-price upper bound on saving.

E.2 Learning fees and fixed-record protocols

The following checks measure total query fees on the fixed protocols.

Whole-batch learning fees. Use the one-hint interface with $(r, c_h, \varepsilon_h) = (10, 1, 0.1)$, $\beta_0 = 0.05$ and $\chi_{\mathrm{rec}} = 100$. For each odd N, designate $s = (N-1)/2$ pilots, s reserve items and one background item. The threshold is $(N+1)/2$. The graph learner receives the support, fees and public direct-result routine. It estimates means from the purchased pilots. True response probabilities are used only to draw mechanisms and compute the oracle comparison. Sixteen independent batch seeds are used at each size. The same seed indices are reused across sizes. Uncertainty estimates use the independent draws within each size. Figure 1b shows all batches and their mean saving with one standard error across the sixteen independent draws.

Let Q_H count purchased hints and let U count results left unread at termination. Every query is charged once, so

$$\begin{aligned} C_{\mathrm{paid}} &= r(N-U) + c_h Q_H, \\ Nr - C_{\mathrm{paid}} &= rU - c_h Q_H. \end{aligned}$$

The 41 initial full-completion batches have $U = Q_H = 0$. On the other branches, net savings are positive exactly when $rU > c_h Q_H$.

For the known-law controller, the sufficient generic reserve at $N = 801$ is

$$\left\lceil 160 \left[\sqrt{801} \sqrt{\log(3 + \sqrt{801}/2)} + \tfrac{4}{3} \log(3 + \sqrt{801}/2) \right] + 161 \right\rceil = 8402.$$

This sufficient reserve exceeds the available supply. The exact dynamic program in Appendix E.1 evaluates its finite optimum.

Immutable code-result replay. Use the frozen initial full-test records from the Qwen-7B and Llama-8B BigCodeBench runs (Zhuo et al., 2025). Select the registered environment-recovery record where one exists. Include every record with a valid execution and a complete, nonduplicated vector of pass/fail/error test statuses. A pass is one; fail and error are zero. Module errors, missing tests and nonbinary statuses are ineligible. This gives 148 Qwen and 144 Llama programs. The four excluded Llama task IDs are 89, 184 and 593 for invalid execution or module error, and 409 for an incomplete or duplicate test list. The archive records the precise reason and source hash for each record.

A legal query discloses one test bit. Tests are ordered by name and programs by numerical task ID. The full-result baseline stops each program at its first zero; it reads all tests of a passing program. For a program with m_g tests, let Q_g be its purchased test addresses. The aggregate rule accepts when at

least k programs pass. Its support endpoints are

$$l_g = \mathbf{1}\{|Q_g| = m_g, \text{ all purchased bits are } 1\},$$
$$u_g = \mathbf{1}\{\text{no purchased bit is } 0\}, \qquad L = \sum_g l_g, \quad U = \sum_g u_g,$$
$$\{L \geq k\} \Longrightarrow K \geq k \quad \text{for every compatible completion,}$$
$$\{U < k\} \Longrightarrow K < k \quad \text{for every compatible completion.}$$

The policy reads the next program to its first zero or complete pass. It updates $[L, U]$ and stops at one of these two certificates. The simulator uses hidden bits only to answer a purchased query. An independent replay checks address uniqueness, fees and terminal bounds.