THE RECOVERABILITY GAP IN COORDINATION WITH PRIVATE OBJECTIVES

Anonymous authors
Paper under double-blind review

ABSTRACT

Accurate local optimization responses need not reveal the objective differences that matter for a shared decision. We define the recoverability gap as the excess expected loss of the best transcript policy relative to the best instance-aware planner under the same protocol. For strongly convex objectives, we identify changes that preserve every allowed local response yet alter the shared optimum. When response maps coincide, the gap is the cost of choosing one common allocation rather than instance-specific optima. It vanishes precisely when the instances share an optimum almost surely. The result covers continuous priors and boundary optima. Constrained sensitivity quantifies the loss from small hidden changes. Local messages can select an optimum within a known family or certify a proposed allocation without a candidate list. Controlled eight-party experiments test the predicted loss and recovery using the same posterior-optimal decision rule. In generated resource-allocation tasks without inserted hidden perturbations, a record search finds decision-relevant ambiguity in short transcripts. Greater tested depths remain unresolved. These results link the differences an interface conceals to the shared decisions it can support.

1 INTRODUCTION

Distributed coordination uses local responses to choose allocations under shared constraints. In energy management, for example, parties respond to prices or messages while a coordinator maintains system-wide balance without observing their objectives (Palomar & Chiang, 2006; Kraning et al., 2014). Communication and local computation limit how many responses it can obtain and how accurately local problems are solved (Jiang et al., 2024; Takezawa et al., 2025). A distinct question remains even with exact local solves: do the responses reveal enough to choose the allocation that would be optimal if the objectives were known?

More queries need not reveal the relevant distinction. Figure 1 illustrates two assignments of private objectives that agree throughout the region reached by the allowed local responses. The shared constraint places the terminal allocation elsewhere, where the assignments require opposite reallocations. Every permitted finite adaptive query sequence then has the same transcript under both instances. This example raises a broader question: which objective differences are invisible to an interface, yet relevant to the decision it must support?

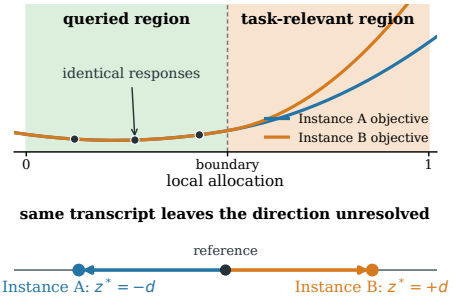


Figure 1: Queries probe where objectives agree. Shared optima lie where they differ.

We compare two information structures under the same protocol, feasibility constraints, and resource budgets. A transcript policy uses the public interaction history to choose queries and actions. An instance-aware planner may also use the private objectives. Both may adapt as responses arrive. The recoverability gap is the excess expected loss of the best transcript policy relative to the best instance-aware planner. It measures which improvements knowledge of the private objectives makes possible under the fixed interface.

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.

Our analysis separates response invisibility from decision relevance. Convex optimality determines whether an objective change preserves every allowed local response and whether it alters the shared optimum. A local active constraint can hide a change, while a terminal active constraint can make it irrelevant to the decision. For strongly convex objectives with common response maps, an exact Bayes identity measures the loss of one common allocation relative to instance-specific optima, allowing continuous priors and boundary optima. The gap is zero exactly when an optimum is shared almost surely. Constrained sensitivity quantifies small-perturbation loss through feasible directions and curvature, using critical-cone quadratic programs when polyhedral constraints are active.

Additional messages address the missing information in two ways. In the binary example, a party computes a bit from its marginal cost at a public allocation to select the instance-aware optimum. Without a candidate list, bounded-error local subgradients and a public curvature bound instead certify the excess loss of a proposed allocation. The first message resolves a decision. The second certifies its quality. Ordinary responses can also suffice: by standard dual optimality, an allowed price that clears the shared constraint returns an optimal allocation. Identical response maps covering such a price therefore cannot conceal different unique optima. Whether the interface covers that price is distinct from the effort needed to find it.

We first test the mechanism on controlled eight-party, 24-hour tasks with temporally coupled objectives and shared balance constraints. Every information condition uses the same terminal decision rule: minimize the posterior expected objective over a public finite family. This avoids differences in predictor training or coordination updates. Numerical solver accuracy is checked separately. The gaps follow the predicted small-perturbation scale. More ordinary queries leave the records indistinguishable, a normal-gradient control leaves the optimum unchanged, and local sign bits progressively recover the instance-aware value. We then examine generated resource-allocation tasks based on public electricity records, without inserting hidden perturbations. Within a prespecified penalty envelope, a search finds objectives that reproduce short response records but require different optimal allocations. At greater depths, the remaining ambiguity is unresolved by the search and available upper bounds. This diagnostic concerns one prescribed trajectory, complementing the controlled test of equality over an entire query domain. The computed marginal-message bounds cover the evaluated losses, but none meets the prescribed stopping tolerances.

2 Related Work

Coordination under limited resources. Decomposition methods coordinate coupled objectives through prices, messages, or consensus variables (Nedić & Ozdaglar, 2009; Boyd et al., 2011). Proximal message passing applies this approach to dynamic energy networks (Kraning et al., 2014). Federated, parallel, and decentralized optimization study the effects of communication, local computation, query depth, and inexact solves on convergence and resource complexity (McMahan et al., 2017; Jiang et al., 2024; Jambulapati et al., 2024; Takezawa et al., 2025). We hold the coordination protocol and resource limits fixed and ask whether its responses support the shared decisions that knowledge of the objectives would permit.

Adaptivity and the value of information. Classical work studies adaptive stochastic allocation and information gathering (Dean et al., 2008; Golovin & Krause, 2011). Feature acquisition, partial-feedback learning, cost-sensitive decisions, and costly sensing examine which observations are worth obtaining (Valancius et al., 2024; Heuillet et al., 2024; Li & Oliva, 2025; Kapoor & Nair, 2026). Uncertainty measures can also be assessed through their role in prediction and data acquisition (Bickford Smith et al., 2025). Blackwell's comparison of experiments evaluates information through the decisions it supports (Blackwell, 1953; Khan et al., 2024). De Lara & Gossner (2020) characterize positive information value using the convex geometry of payoffs, beliefs, and optimal actions. We study how such information is generated by local optimization queries: which objective changes preserve their response maps, and which of these changes alter the shared optimum. Both policy classes may adapt. Only the instance-aware planner observes the objectives.

Indistinguishability and partial observations. Two-point testing yields lower bounds from indistinguishability (Le Cam, 1986), including when observations are chosen interactively (Chen et al., 2024). Our two-action bound uses this argument. The perturbation result combines convex optimality and constrained sensitivity (Boyd & Vandenberghe, 2004; Shapiro, 1985) to identify response-invisible changes, characterize when they affect a shared optimum, and quantify the induced loss. Partial

monitoring and partially observable control also distinguish hidden information from the feedback available to a policy (Bartók et al., 2014; Kaelbling et al., 1998). Related learning studies address policy-induced distribution shift (Ross et al., 2011), representations and memory under partial observations (Zhang et al., 2024; Eberhard et al., 2025), and privileged information during training (Cai et al., 2024; Huang et al., 2025; Song et al., 2025). We distinguish the best policy allowed by an information structure from the performance of a fitted predictor using that information. Inverse optimization estimates objectives compatible with observed decisions (Bertsimas et al., 2015). Our record diagnostic instead measures how far the optimal shared allocations of compatible objectives can differ.

3 Private-Objective Coordination and the Recoverability Gap

The model separates the coordinator's information from its permitted decisions. Sections 4–5 focus on terminal allocations chosen after local queries. The general action model also covers changes to a query schedule, used in the query-scheduling experiments. Appendix A specifies the full information structure and feasibility conditions.

3.1 Private Instances and Online Transcripts

A coordination instance is $I = (I_{\mathrm{pub}}, \theta)$, with $\theta \sim \nu$. The public protocol I_{pub} specifies the horizon H, initial-state law and dynamics, query and action spaces, shared constraints, reference policy, and resource limits. The private instance $\theta \in \Theta$ determines the local objectives, response kernels $O_{\theta,t}$, and terminal loss ℓ_θ. Both policy classes are evaluated under the same public prior ν.

At stage $t = 0, \dots, H-1$, let S_t be the public coordination state and τ_t^- the pre-query transcript. The coordinator chooses q_t and observes $Y_t \sim O_{\theta,t}(\cdot \mid \tau_t^-, q_t)$, giving $\tau_t^+ = (\tau_t^-, q_t, Y_t)$. It then chooses $A_t \sim \pi_t(\cdot \mid \tau_t^+)$, and the public state evolves as $S_{t+1} \sim P_t(\cdot \mid S_t, q_t, Y_t, A_t)$. The transition kernel P_t is common to all private instances. The transcript records public states, queries, responses, actions, and remaining resources. For a transcript policy, private objectives affect the observed record through the response kernels.

3.2 Transcript Policies and Instance-Aware Planners

A transcript policy uses $q_t \sim \mu_t(\cdot \mid \tau_t^-)$ and $A_t \sim \pi_t(\cdot \mid \tau_t^+)$. Let Π_{tr} be the class of these policies. An instance-aware planner follows the same protocol and feasibility constraints but may also condition on θ: $q_t \sim \mu_t^{\mathrm{full}}(\cdot \mid \theta, \tau_t^-)$ and $A_t \sim \pi_t^{\mathrm{full}}(\cdot \mid \theta, \tau_t^+)$. Its policy class is Π_{full}. Since a planner may ignore θ, $\Pi_{\mathrm{tr}} \subseteq \Pi_{\mathrm{full}}$.

3.3 Optimal Values and the Recoverability Gap

We compare the best expected losses under common feasibility and resource limits. The protocol fixes a horizon H, query limit Q, and budget B, and may also limit departures from a deterministic admissible reference policy π_{ref}. Write K for this intervention quota and suppress H, Q in the value notation. At the current post-response transcript, define $A_t^{\mathrm{ref}} = \pi_{\mathrm{ref},t}(\tau_t^+)$ and $Z_t^\pi = \mathbf{1}\{A_t^\pi \neq A_t^{\mathrm{ref}}\}$. A policy is K-sparse if $\sum_{t=0}^{H-1} Z_t^\pi \leq K$ almost surely. Setting $K = H$ leaves this constraint inactive. The reference action is evaluated at the transcript actually reached, including previous interventions.

Let $\mathcal{F}$ be the public terminal feasible set. Write $\Pi^{(K)}_{\mathrm{tr,feas}}(B)$ and $\Pi^{(K)}_{\mathrm{full,feas}}(B)$ for the budget-feasible, K-sparse policies in the two classes satisfying $S_H^\pi \in \mathcal{F}$ almost surely. For a policy π, let $L_\theta(\pi) = \mathbb{E}_\theta^\pi[\ell_\theta(S_H)]$, where the expectation is over the interaction randomness. The optimal Bayes losses are

$$V^{(K)}_{\mathrm{full}}(\nu, B) = \inf_{\pi \in \Pi^{(K)}_{\mathrm{full,feas}}(B)} \mathbb{E}_{\theta\sim\nu} L_\theta(\pi), \qquad V^{(K)}_{\mathrm{tr}}(\nu, B) = \inf_{\pi \in \Pi^{(K)}_{\mathrm{tr,feas}}(B)} \mathbb{E}_{\theta\sim\nu} L_\theta(\pi). \tag{1}$$

The infimum defining $V^{(K)}_{\mathrm{tr}}$ is over a single policy shared across instances. It adapts to the realized instance through the transcript. The instance-aware value $V^{(K)}_{\mathrm{full}}$ is the oracle benchmark.

Definition 3.1 (Recoverability gap)**.** At budget B and intervention quota K, define

$$\mathrm{Gap}^{(K)}_{\mathrm{rec}}(\nu, B) = V^{(K)}_{\mathrm{tr}}(\nu, B) - V^{(K)}_{\mathrm{full}}(\nu, B). \tag{2}$$

The two values share all protocol constraints and differ only in access to θ. Their gap is nonnegative by policy-class inclusion. Let $V_{\mathrm{ref}}(\nu, B)$ be the reference Bayes loss. Omitting (ν, B),

$$\mathrm{Gap}_{\mathrm{rec}}^{(K)} = \big(V_{\mathrm{ref}} - V_{\mathrm{full}}^{(K)}\big) - \big(V_{\mathrm{ref}} - V_{\mathrm{tr}}^{(K)}\big). \tag{3}$$

The gap is the part of the instance-aware improvement that the best transcript policy cannot recover.

3.4 Transcript Aliasing

Two private instances are *transcript equivalent* if every adaptive, randomized transcript policy admissible under both induces the same law for each transcript prefix. Write $P_{\theta,t}^{\pi,-}$ and $P_{\theta,t}^{\pi,+}$ for the laws of τ_t^- and τ_t^+. We denote this equivalence by $\theta_0 \equiv_{\Pi_{\mathrm{tr}},H} \theta_1$: for every such π, $P_{\theta_0,t}^{\pi,-} = P_{\theta_1,t}^{\pi,-}$ for $t = 0, \dots, H$ and $P_{\theta_0,t}^{\pi,+} = P_{\theta_1,t}^{\pi,+}$ for $t = 0, \dots, H-1$. We call this *transcript aliasing*. It is relative to the fixed response interface and public protocol. Appendix A.7 gives the approximate definition in total variation. If aliased instances require different actions, even an optimal transcript policy may incur excess loss.

4 Identical Responses, Different Optimal Allocations

We begin with an explicit two-party example, then identify a general condition for query-invisible objective changes to alter a shared optimum. The classical two-point testing bound and its approximate versions are given in Appendix B.

4.1 A Strongly Convex Coordination Problem

The allowed price queries return local optima in a region where the private objectives agree, while the shared constraint requires final allocations in a region where they differ. For $\kappa > 0$, define $\varphi_\kappa : [0,1] \to \mathbb{R}$ by

$$\varphi_\kappa(x) = \begin{cases} \dfrac{1}{2}\left(x - \dfrac{1}{4}\right)^2, & 0 \le x \le \dfrac{1}{2}, \\ \dfrac{1}{32} + \dfrac{1}{4}\left(x - \dfrac{1}{2}\right) + \dfrac{\kappa}{2}\left(x - \dfrac{1}{2}\right)^2, & \dfrac{1}{2} < x \le 1. \end{cases} \tag{4}$$

The branches match in value and derivative at $x = 1/2$. Thus φ_κ has a $\max\{1,\kappa\}$-Lipschitz derivative and is $\min\{1,\kappa\}$-strongly convex. Fix public constants $0 < \kappa_L < \kappa_H$. Consider two equiprobable private instances with the following objective assignments:

$$\theta_0 : (f_1, f_2) = (\varphi_{\kappa_L}, \varphi_{\kappa_H}), \qquad \theta_1 : (f_1, f_2) = (\varphi_{\kappa_H}, \varphi_{\kappa_L}). \tag{5}$$

After at most Q price queries, the coordinator chooses a terminal allocation in

$$\mathcal{X} = \{x \in [0,1]^2 : x_1 + x_2 = 3/2\}, \qquad x^{\mathrm{ref}} = (3/4, 3/4), \tag{6}$$

with loss $\ell_\theta(x) = f_{1,\theta}(x_1) + f_{2,\theta}(x_2)$. The intervention quota is $K = 1$. A query to party $m \in \{1,2\}$ at price $\lambda \in [-1/4, 1/4]$ returns only

$$\mathcal{R}_{m,\theta}(\lambda) = \arg\min_{x \in [0,1]} \{f_{m,\theta}(x) + \lambda x\}. \tag{7}$$

Theorem 4.1 (Continuous convex recoverability gap)**.** *For every finite query limit Q, every adaptive, randomized, party-specific query policy in the public price domain receives*

$$\mathcal{R}_{m,\theta}(\lambda) = \frac{1}{4} - \lambda \tag{8}$$

in both instances. Let $\Sigma = \kappa_L + \kappa_H$, $d = (\kappa_H - \kappa_L)/(4\Sigma)$. The unique instance-aware optima are

$$x^{\star,0} = (3/4 + d, 3/4 - d), \qquad x^{\star,1} = (3/4 - d, 3/4 + d). \tag{9}$$

Under the uniform prior, the recoverability gap is

$$G = \mathrm{Gap}_{\mathrm{rec}}^{(1)} = \frac{(\kappa_H - \kappa_L)^2}{32(\kappa_H + \kappa_L)}. \tag{10}$$

The symmetric reference attains the optimal transcript value. Among transcript policies, the minimax excess loss relative to the instance-aware optima also equals G. For $(\kappa_L, \kappa_H) = (1, 3)$, $G = 1/32$.

All returned local optima lie in $[0, 1/2]$, where the objectives agree. Thus the same adaptive query policy induces identical transcript laws under the two instances. The shared constraint places each terminal coordinate in $[1/2, 1]$, where the curvature assignment determines the coordinated optimum. To compute the gap, write $x = (3/4+z, 3/4-z)$, with $z \in [-1/4, 1/4]$, and let $\ell^\star_{\theta_i} = \ell_{\theta_i}(x^{\star,i})$. The two excess losses are $\ell_{\theta_0}(x) - \ell^\star_{\theta_0} = \frac{\Sigma}{2}(z-d)^2$ and $\ell_{\theta_1}(x) - \ell^\star_{\theta_1} = \frac{\Sigma}{2}(z+d)^2$. For any transcript policy, the terminal displacement Z has the same law under both instances. The policy's Bayes excess loss is therefore $\frac{\Sigma}{2}\big(\mathbb{E}[Z^2] + d^2\big) \geq \frac{\Sigma d^2}{2} = G$, with equality at the reference $Z = 0$. Appendix C gives the full proof.

What the interface reveals. At a returned point $x = 1/4 - \lambda$, the objective value $\lambda^2/2$ and derivative $-\lambda$ are independent of the private curvature. Returning either quantity therefore preserves aliasing. At interior returned points the second derivative is also 1. At the splice point $x = 1/2$ it need not exist. In contrast, a derivative query at the reference coordinate $3/4$ distinguishes the assignments, as does an exact price query at $\lambda^\star = -1/4 - \kappa_L/4$. Appendix C.4 gives the precise interface variants. The obstruction concerns which part of the objective the interface probes, not the accuracy of the local solve. The normalized interval is not special. For any $L > 0$, scaling allocations by $4L$ and objectives by $(4L)^2$ produces the same separation for the prescribed price domain $[-L, L]$. Hence no finite query count resolves the pair through that bounded interface, although a query outside its blind range or a decision-relevant local derivative does. Appendix C.4 states and proves this scale-invariant form.

4.2 Common Responses and Decision-Relevant Differences

Consider a terminal allocation in a nonempty compact convex set $\mathcal{X} = \{x \in \prod_m D_m : Ax = b\}$. Both policy classes may choose any $x \in \mathcal{X}$, including without querying. The private objective is $F_\theta(x) = \sum_m f_{m,\theta}(x_m)$, and $\theta \sim \nu$ may have continuous support. An ordinary query returns

$$R_{m,\theta}(q) = \arg\min_{u \in D_m} \{f_{m,\theta}(u) + \psi_m(u; q)\}, \qquad q \in \mathcal{Q}_m, \tag{11}$$

where the public convex term ψ_m and its full parameter range specify the interface.

Theorem 4.2 (Common responses and decision loss)**.** *Suppose F_θ is a measurable family of continuous μ-strongly convex functions on $\mathcal{X}$, with $\mathbb{E}_\nu \sup_{x\in\mathcal{X}} |F_\theta(x)| < \infty$. On one common full-measure set of instances, assume all permitted response maps are identical. Write $\overline{F} = \mathbb{E}_\nu F_\theta$, $\overline{x} = \arg\min_{\mathcal{X}} \overline{F}$, and $x^\star_\theta = \arg\min_{\mathcal{X}} F_\theta$. Every finite adaptive, randomized transcript policy then has a terminal-output law independent of θ, and*

$$G_\nu = \overline{F}(\overline{x}) - \mathbb{E}_\nu F_\theta(x^\star_\theta) = \mathbb{E}_\nu[F_\theta(\overline{x}) - F_\theta(x^\star_\theta)] \geq \frac{\mu}{2}\mathbb{E}_\nu\|\overline{x} - x^\star_\theta\|^2. \tag{12}$$

In particular, $G_\nu = 0$ exactly when $x^\star_\theta = \overline{x}$ almost surely.

Proof idea. With a common random tape, identical histories give the same next query and response, so the record leaves the prior unchanged. Its best terminal decision therefore minimizes $\overline{F}$. Randomization cannot lower that value. The instance-aware planner minimizes each F_θ separately. Strong convexity lower-bounds this excess loss by the mean squared distance from the common allocation to the instance-specific optima. This argument allows boundary optima and does not require a finite list of objectives. Appendix L.1 gives the measurability details and a minimax lower bound for any objective class containing an aliased pair. Computing the Bayes-optimal decision still requires its stated prior model.

Checking local invisibility. For differentiable convex F, optimality over a convex set C is equivalent to $-\nabla F(x) \in N_C(x)$, where $N_C(x) = \{n : \langle n, y - x\rangle \leq 0 \ \forall y \in C\}$ (Bertsekas, 2012). Take $f_{m,\sigma} = f_{m,0} + \sigma\alpha h_m$, $\sigma = \pm 1$, on compact convex D_m. Assume $f_{m,0}, h_m, \psi_m(\cdot; q)$ are differentiable near D_m, both $f_{m,\sigma}$ are strongly convex on D_m, and each query term is convex (Appendix L.2). Let $R_{m,0}$ be the response without the perturbation. By uniqueness of the local optimizers, both responses equal $r = R_{m,0}(q)$ exactly when

$$-\nabla f_{m,0}(r) - \nabla_u \psi_m(r; q) - \sigma\alpha\nabla h_m(r) \in N_{D_m}(r) \qquad (\sigma = \pm 1). \tag{13}$$

A simple sufficient condition is $\nabla h_m(R_{m,0}(q)) = 0$ for every $m,\ q \in \mathcal{Q}_m$. At a local boundary, Eq. (13) can hold even with a nonzero perturbation gradient: the active constraint absorbs the change. Appendix L.6 gives a continuous-parameter quadratic example of this saturation mechanism.

When a hidden change affects the decision. Response invariance alone does not imply decision loss. We next ask whether the same allocation remains optimal after the hidden change.

Corollary 4.3 (Perturbations and constrained geometry)**.** *Let $F_\sigma = F_0 + \sigma\alpha h$, with $\alpha > 0$, be differentiable and μ-strongly convex, with identical ordinary response maps and equiprobable signs. For $x_0^\star = \arg\min_{\mathcal{X}} F_0$ and $x_\sigma^\star = \arg\min_{\mathcal{X}} F_\sigma$,*

$$G(\alpha) = F_0(x_0^\star) - \frac{1}{2}\sum_{\sigma=\pm 1} F_\sigma(x_\sigma^\star) \geq \frac{\mu}{8}\|x_+^\star - x_-^\star\|^2. \tag{14}$$

The gap is zero exactly when $-\nabla F_0(x_0^\star) - \sigma\alpha\nabla h(x_0^\star) \in N_{\mathcal{X}}(x_0^\star)$ for both signs. If $x_0^\star$ is interior to the local domains, let Z have orthonormal columns spanning $\ker A$. Then $G(\alpha) > 0$ exactly when $g = Z^\top \nabla h(x_0^\star) \neq 0$. Under this interior condition, if the objectives are C^2 and uniformly strongly convex for all sufficiently small signed amplitudes, then

$$G(\alpha) = \frac{\alpha^2}{2} g^\top H_T^{-1} g + o(\alpha^2), \qquad H_T = Z^\top \nabla^2 F_0(x_0^\star) Z. \tag{15}$$

The normal-cone criterion distinguishes two roles of constraints: a local boundary can hide an objective difference, while a terminal boundary can make that difference irrelevant to the decision. The interior formula is its especially simple case. For active polyhedral constraints, Appendix L.5 replaces the inverse-Hessian expression by two quadratic programs over the critical cone. Appendix L.7 constructs smooth invisible perturbations, and Appendix L.8 extends the expansion to bounded continuous parameter laws under uniform remainder control.

When the ordinary interface is sufficient. For price responses with query term $\lambda^\top A_m x_m$, any allowed price whose local responses satisfy $\sum_m A_m R_m(\lambda) = b$ produces a global optimum by standard dual optimality (Boyd & Vandenberghe, 2004). Two identical response maps containing such a clearing price cannot conceal different unique optima. Appendix L.9 separates this coverage condition from the query complexity of finding that price. Appendix D treats the two-party example when the allocation is reached through bounded increments. It separates the effect of a larger reachable set from the information supplied by the interface.

5 Decision-Relevant Certificates

When ordinary responses leave a decision unresolved, an additional local message may supply the missing distinction. We first show how one bit selects the optimal allocation in the two-party construction. We then use local subgradients to certify the quality of a proposed allocation without identifying the objective.

For the one-party scheme below, let W be the ordinary record after the last response and before the coordination action. A decision certificate is a message C computed by a designated party from its local private information and W, possibly with independent local randomness. The coordinator acts on (W, C). Certificate communication has a separate bit budget B_C. The ordinary query domain and resource limits are unchanged. Let $V^{(1)}_{\mathrm{tr}+C}(\nu; B, B_C)$ be the optimal Bayes loss of this augmented class with one intervention and the same terminal feasibility constraint. Appendix F formalizes the class and proves the results below.

5.1 A Locally Computed One-Bit Certificate

In the two-party construction of Section 4.1, the optimal displacement has public magnitude d but an unknown direction. The protocol specifies the two curvature values, κ_L and κ_H, with one assigned to each party. At the symmetric reference, the two possible marginal costs are

$$m_L = \frac{1}{4} + \frac{\kappa_L}{4}, \qquad m_H = \frac{1}{4} + \frac{\kappa_H}{4}. \tag{16}$$

Party 1 sends the locally computed bit

$$C_1 = \mathbf{1}\{f_1'(3/4) = m_L\}. \tag{17}$$

Since $m_L < m_H$, $C_1 = 1$ means that a small shift of allocation from party 2 to party 1 lowers the loss at the reference.

Theorem 5.1 (One-bit certificate sufficiency)**.** *For the hard pair of Theorem 4.1, the bit in Eq.* (17) *gives the allocation* $x_1 = \frac{3}{4} + (2C_1 - 1)d$, $x_2 = \frac{3}{2} - x_1$. *This is the unique instance-aware optimum under either instance. Under the uniform prior,* $V^{(1)}_{\mathrm{tr}+C}(\nu; B, 1) = V^{(1)}_{\mathrm{full}}(\nu, B)$.

Party 1 sends the bit once, immediately before the terminal allocation, using $B_C = 1$.

How much information is needed? The augmented record must separate instances with different uniquely optimal allocations. For the pair above, a deterministic fixed-length certificate closing the gap needs at least one bit. For a finite aliased family with N distinct optimal actions, it needs at least $\lceil \log_2 N \rceil$ bits. This is attainable when a designated party can identify the action class locally. If the bit above is independently flipped with known probability $p \le 1/2$, the optimal displacement shrinks by $1 - 2p$ and the remaining gap is $4p(1-p)G$. Appendix F proves these necessity and noise statements. The perturbation family admits the same local rule when a party's perturbation gradient is nonzero at a public point: its marginal identifies the hidden sign. Multiple signs may require several parties' messages. The experiments use this rule with a common posterior-optimal decoder. Appendix L.8 gives the finite-family value formula.

5.2 Approximate Certification without a Candidate List

Without a known candidate list, local messages can serve a different purpose: certifying a proposed allocation rather than selecting it from a finite set. Section 4.2 used constrained optimality to determine whether a hidden change moves the optimum. Here we use additional local messages to bound how far a proposed allocation is from satisfying that condition. Let $x \in \mathcal{X}$ be public. Each party computes a local subgradient at x_m and reports an approximation with a certified error bound. The resulting vector $\widehat{g}$ approximates some $g \in \partial F(x)$ with $\|\widehat{g} - g\| \le \delta$. For smooth objectives, $g = \nabla F(x)$.

Proposition 5.2 (List-free optimality certificate)**.** *If* F *is* μ*-strongly convex on the common convex feasible set and* $x \in \mathcal{X}$*, define* $\widehat{r}(x) = \mathrm{dist}(-\widehat{g}, N_{\mathcal{X}}(x))$. *Then*

$$F(x) - F^\star \le \frac{(\widehat{r}(x) + \delta)^2}{2\mu}. \tag{18}$$

The coordinator needs the public constraints, a valid curvature lower bound, and the received subgradients and error bounds, but no prior or list of candidate objectives. The bound follows by adding a normal vector to the strong-convexity inequality and completing the square. Appendix F.6 gives the proof and message accounting. Section 6.2 evaluates it at allocations of a fixed coordination protocol, counting the additional local solves and messages separately from ordinary responses.

6 Experiments

The experiments address two questions: does the predicted information loss occur when the mechanism's conditions hold, and can finite response records conceal decision-relevant differences in tasks without inserted hidden perturbations? We use controlled perturbation tasks for the first question and generated resource-allocation tasks for the second. Both use public hourly electricity records from Open Power System Data (OPSD), version 2020-10-06 (Open Power System Data, 2020). Their private objectives are generated, not observed institutional costs. Additional studies examine the behavior of specific response protocols and query-scheduling rules.

6.1 A Direct Test of the Perturbation Mechanism

Each of eight parties chooses a 24-hour allocation. Public hourly balance constraints couple the parties, and non-diagonal local Hessians couple adjacent hours. We form $F_\pm = F_0 \pm \alpha h$ with a perturbation whose gradient vanishes at every response allowed by the prescribed price domain. The same construction yields a control with perturbation gradient normal to the balance constraint at the baseline optimum. All terminal optima evaluated here lie in the perturbation's quadratic branch. We use four public load days, five objective seeds, three temporal coupling strengths, and seven amplitudes, giving 420 pairs per perturbation type. Appendix M specifies the task, full query domain, and parameter grid.

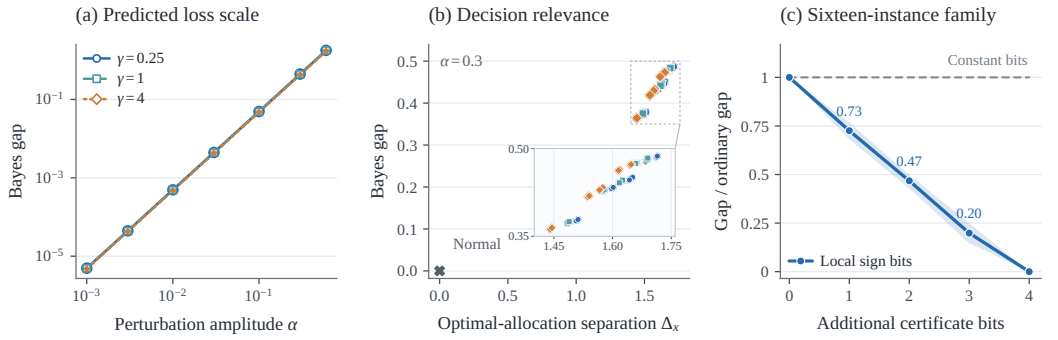


Figure 2: Direct mechanism tests. (a) Mean Bayes gaps (markers) and small-perturbation predictions (lines). (b) Decision relevance at $\alpha = 0.3$. The inset magnifies the relevant cases. (c) Mean remaining gap normalized by the mean ordinary gap. Shading spans per-task ratios.

Table 1: Common-decoder comparison on 60 tasks at $\alpha = 0.3$. The gap is the mean raw Bayes excess loss. Ordinary contacts return 24 float64 values each. Informative and constant bits use the same local derivative evaluation. The expanded query changes the allowed price domain.

Information available	Additional response	Remaining gap
Ordinary, $Q = 0, 8, 32, 128$	None	0.438435
Ordinary + constant message	1 bit	0.438435
Ordinary + local sign	1 bit	$< 10^{-12}$
Ordinary + expanded-price query	24 float64 values	$< 10^{-12}$
Instance-aware	Private sign known	0

The finite objective family and its prior are public. Every information condition uses the same decoder, which minimizes the posterior expected objective over the common feasible set, with the same convex solver and tolerance. Only the available information changes. Appendix M.5 reports numerical accuracy checks. We report unnormalized Bayes excess loss, so the cancellation in Eq. (14) applies directly.

Invisibility and loss scale. Ordinary transcripts are bit-for-bit identical under both signs at query depths $Q \in \{0, 8, 32, 128\}$. Appendix M.2 proves equality over the entire permitted price domain. Every decision-relevant pair has positive gap. At $\alpha = 0.3$, the mean is 0.438435 and the optimal-allocation separation ranges from 1.440 to 1.715. Across $\alpha \leq 0.01$, the largest relative discrepancy between the computed gap and the leading term in Eq. (15) is 4.85×10^{-6}, without fitting a scale factor. The normal perturbations have zero allocation separation and gap at most 3.56×10^{-15} in magnitude (Figure 2a,b). These controls separate response invisibility from its effect on feasible reallocations.

Recovering the missing distinction. Increasing ordinary query depth or adding a constant bit leaves the posterior and Bayes gap unchanged. A bit computed from one party's local marginal identifies the binary sign. The same decoder then reaches the instance-aware value. One additional price outside the ordinary domain also distinguishes the pair. Table 1 separates these changes in message content and query permission. With four independent hidden signs, the 16-instance family requires a richer certificate: the mean remaining gaps after zero through four local sign bits are 0.438642, 0.318287, 0.205159, 0.087085, and zero. All signs enter the same balance-constrained problem with temporally coupled local objectives.

6.2 Decision Ambiguity in Resource-Allocation Records

Here each of eight parties holds resources with capacity bounds, preferred allocations, and quadratic plus absolute-deviation costs. Public hourly balance couples their aggregate allocations. Appendix G specifies this resource-cost model. Unlike the controlled tasks, these objectives contain no inserted

hidden perturbation. We fix four additional development dates, five objective seeds, and two conflict levels before evaluation, giving 40 tasks.

For each ordinary response record, we search a prespecified penalty envelope for objectives that reproduce every transmitted message but require different optimal allocations. Preferences, capacities, curvatures, and demand remain fixed. Matching uses the actual 50-step solver and the declared encoding: either float64 messages or a fixed-point grid of width 2^{-12}. A verified pair gives a lower bound on the record-conditioned worst-case decision loss. This asks what remains possible after one prescribed trajectory, whereas Section 6.1 tests equality over the full query domain. Appendix N defines the loss and reports all search outcomes. After the eight initialization contacts, the lower bound exceeds 10^{-4} in 22 of 40 float64 records and all 40 fixed-point records. After eight further contacts, the counts are zero and 40, respectively. After 32 or 128 further contacts, neither format yields a bound above the threshold. These later cases remain unresolved: the candidate search is incomplete, and the available upper bounds exceed the threshold. The findings establish ambiguity in short records within the specified envelope, but do not establish recovery from the longer records.

Certifying the proposed allocations. We next request additional local marginal messages at each of the 320 task–format–depth allocations and apply Proposition 5.2. Two marginal-message precisions give 640 certificate evaluations. Every computed bound covers the evaluated excess loss, but none is at most either prescribed tolerance, 10^{-3} or 10^{-4}. Thus the bounds do not provide a stopping certificate at these accuracies in this task. Appendix N.4 reports their sizes and additional communication, and discusses the conservative public curvature bound and the selected nonsmooth subgradients. These messages assess the proposed allocation. They are not part of the preceding ordinary record.

6.3 Additional Protocol and Query-Scheduling Studies

Using the same resource-cost model on separate evaluation sets, these studies examine specific protocols and query-scheduling rules, rather than optimal transcript policies.

Response protocols. At matched contacts, message dimensions, local bisection work, and projection counts, the aggregate-optimizer and marginal-value protocols attain mean normalized regrets of 1.219×10^{-3} and 6.264×10^{-3} after 32 post-initialization contacts. Their paired difference is 5.045×10^{-3} (95% date-block interval $[4.393, 5.740] \times 10^{-3}$). Both query points and public updates differ. Appendix K details the protocols, tuning, and subgradient information implicit in an exact aggregate response.

Query scheduling. Across 2,400 runs, three instance-aware query replacements capture 98.70% of the improvement found by the same sequential evaluator without a quota, changing 11.84% of reference positions and reducing regret by 9.10%. Exact joint search on 480 separate runs also benefits from multiple replacements. Appendix H gives the search classes, uncertainty, and pre-projection violation checks. When an evaluator supplies alternatives, an online residual-based rule captures 26.46% of the exact three-intervention evaluator's improvement at a 5.99% trigger rate. Raw dual changes give similar capture with more triggers. The paired capture-difference interval includes zero. On fixed records, net capture is selected signed advantage divided by total positive advantage. At matched selection rates, its mean falls from 35.8% within budget to 1.6% when that budget is excluded from fitting and model selection. Appendix J separates offline ranking from online execution and reports the timing and transfer comparisons. Appendix I studies reference anchoring.

7 Conclusion

Local responses can conceal objective differences that change a shared optimum. For strongly convex objectives with common response maps, the recoverability gap is the cost of one common allocation relative to instance-specific optima. Convex optimality separates local invisibility from terminal relevance, while constrained sensitivity quantifies small-perturbation loss. Local messages can select an optimum within a known family or certify a proposed allocation without a candidate list. Controlled tests verify the mechanism. Resource-allocation records reveal ambiguity within the specified envelope, not population-wide prevalence. Deeper records remain unresolved. The computed certificates cover the evaluated losses but do not meet the prescribed accuracy. Response interfaces should therefore be assessed by the decisions they distinguish, not only by local solution accuracy.

AI USE STATEMENT

Generative AI tools were used solely for language editing. The authors verified the final manuscript and take full responsibility for the submission.

ETHICS STATEMENT

All experiments run entirely in simulation on synthetic navigation and a standard control benchmark, involving only generated states and trajectories.

REPRODUCIBILITY STATEMENT

The appendices provide the proofs, task constructions, date partitions, query costs, and evaluation measures. Appendix G identifies the public data source and describes preprocessing and task generation. Appendix M specifies the controlled perturbation tests and numerical checks. Appendix N describes the original-task record search, message formats, certificate calculations, and reported results.

REFERENCES

Anastasios N. Angelopoulos, Stephen Bates, Emmanuel J. Candès, Michael I. Jordan, and Lihua Lei. Learn then Test: Calibrating predictive algorithms to achieve risk control. *The Annals of Applied Statistics*, 19(2):1641–1662, 2025. doi: 10.1214/24-AOAS1998.

Gábor Bartók, Dean P. Foster, Dávid Pál, Alexander Rakhlin, and Csaba Szepesvári. Partial monitoring—classification, regret bounds, and algorithms. *Mathematics of Operations Research*, 39(4):967–997, 2014. doi: 10.1287/moor.2014.0663.

Dimitri P. Bertsekas. Convex analysis and optimization: Lecture 12. MIT OpenCourseWare, 6.253, Spring 2012, 2012. URL https://ocw.mit.edu/courses/6-253-convex-analysis-and-optimization-spring-2012/.

Dimitris Bertsimas, Vishal Gupta, and Ioannis Ch. Paschalidis. Data-driven estimation in equilibrium using inverse optimization. *Mathematical Programming*, 153:595–633, 2015. doi: 10.1007/s10107-014-0819-4.

Federico Bianchi, Edoardo Zorzi, Alberto Castellini, Thiago D. Simão, Matthijs T. J. Spaan, and Alessandro Farinelli. Scalable safe policy improvement for factored multi-agent MDPs. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 3952–3973. PMLR, 2024. URL https://proceedings.mlr.press/v235/bianchi24b.html.

Freddie Bickford Smith, Jannik Kossen, Eleanor Trollope, Mark Van Der Wilk, Adam Foster, and Tom Rainforth. Rethinking aleatoric and epistemic uncertainty. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 4345–4359. PMLR, 2025. URL https://proceedings.mlr.press/v267/bickford-smith25a.html.

David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24(2):265–272, 1953. doi: 10.1214/aoms/1177729032.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. URL https://web.stanford.edu/~boyd/cvxbook/.

Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011. doi: 10.1561/2200000016.

Yang Cai, Xiangyu Liu, Argyris Oikonomou, and Kaiqing Zhang. Provable partially observable reinforcement learning with privileged information. In *Advances in Neural Information Processing Systems*, volume 37, pp. 63790–63857, 2024. doi: 10.52202/079017-2038.

Fan Chen, Dylan J. Foster, Yanjun Han, Jian Qian, Alexander Rakhlin, and Yunbei Xu. Assouad, Fano, and Le Cam with interaction: A unifying lower bound framework and characterization for bandit learnability. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-2407.

Michel De Lara and Olivier Gossner. Payoffs-beliefs duality and the value of information. *SIAM Journal on Optimization*, 30(1):464–489, 2020. doi: 10.1137/18M1230049.

Brian C. Dean, Michel X. Goemans, and Jan Vondrák. Approximating the stochastic knapsack problem: The benefit of adaptivity. *Mathematics of Operations Research*, 33(4):945–964, 2008. doi: 10.1287/moor.1080.0330.

Onno Eberhard, Michael Muehlebach, and Claire Vernade. Partially observable reinforcement learning with memory traces. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 14934–14949. PMLR, 2025. URL https://proceedings.mlr.press/v267/eberhard25a.html.

Daniel Golovin and Andreas Krause. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *Journal of Artificial Intelligence Research*, 42:427–486, 2011. doi: 10.1613/jair.3278.

Maxime Heuillet, Ola Ahmad, and Audrey Durand. Neural active learning meets the partial monitoring framework. In *Proceedings of the 40th Conference on Uncertainty in Artificial Intelligence*, volume 244 of *Proceedings of Machine Learning Research*, pp. 1621–1639. PMLR, 2024. URL https://proceedings.mlr.press/v244/heuillet24a.html.

Dongchi Huang, Jiaqi Wang, Yang Li, Chunhe Xia, Tianle Zhang, and Kaige Zhang. PIGDreamer: Privileged information guided world models for safe partially observable reinforcement learning. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 25859–25875. PMLR, 2025. URL https://proceedings.mlr.press/v267/huang25ai.html.

Arun Jambulapati, Aaron Sidford, and Kevin Tian. Closing the computational-query depth gap in parallel stochastic convex optimization. In *Proceedings of the 37th Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pp. 2608–2643. PMLR, 2024. URL https://proceedings.mlr.press/v247/jambulapati24b.html.

Xiaowen Jiang, Anton Rodomanov, and Sebastian U. Stich. Federated optimization with doubly regularized drift correction. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 21912–21945. PMLR, 2024. URL https://proceedings.mlr.press/v235/jiang24e.html.

Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2):99–134, 1998. doi: 10.1016/S0004-3702(98)00023-X.

Vansh Kapoor and Jayakrishnan Nair. MDPs with a state sensing cost. In *Proceedings of the 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, pp. 1648–1656. PMLR, 2026. URL https://proceedings.mlr.press/v300/kapoor26a.html.

M. Ali Khan, Haomiao Yu, and Zhixiang Zhang. On comparisons of information structures with infinite states. *Journal of Economic Theory*, 218:105841, 2024. doi: 10.1016/j.jet.2024.105841.

Matt Kraning, Eric Chu, Javad Lavaei, and Stephen Boyd. Dynamic network energy management via proximal message passing. *Foundations and Trends in Optimization*, 1(2):70–122, 2014.

Lucien Le Cam. *Asymptotic Methods in Statistical Decision Theory*. Springer Series in Statistics. Springer, New York, 1986. doi: 10.1007/978-1-4612-4946-7.

Yang Li and Junier Oliva. Towards cost sensitive decision making. In *Proceedings of the 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 3601–3609. PMLR, 2025. URL https://proceedings.mlr.press/v258/li25h.html.

H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pp. 1273–1282. PMLR, 2017.

Angelia Nedić and Asuman Ozdaglar. Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54(1):48–61, 2009. doi: 10.1109/TAC.2008.2009515.

Open Power System Data. Data package time series. Version 2020-10-06, 2020. URL https://doi.org/10.25832/time_series/2020-10-06. Primary records from the ENTSO-E Transparency Platform.

Daniel P. Palomar and Mung Chiang. A tutorial on decomposition methods for network utility maximization. *IEEE Journal on Selected Areas in Communications*, 24(8):1439–1451, 2006. doi: 10.1109/JSAC.2006.879350.

Stephane Ross, Geoffrey J. Gordon, and J. Andrew Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics*, volume 15 of *Proceedings of Machine Learning Research*, pp. 627–635. PMLR, 2011.

Alexander Shapiro. Second order sensitivity analysis and asymptotic theory of parametrized nonlinear programs. *Mathematical Programming*, 33:280–299, 1985. doi: 10.1007/BF01584378.

Yuda Song, Dhruv Rohatgi, Aarti Singh, and J. Andrew Bagnell. To distill or decide? understanding the algorithmic trade-off in partially observable rl. In *Advances in Neural Information Processing Systems*, volume 38, pp. 89559–89606, 2025. doi: 10.52202/085713-2996.

Yuki Takezawa, Xiaowen Jiang, Anton Rodomanov, and Sebastian U. Stich. Exploiting similarity for computation and communication-efficient decentralized optimization. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 58359–58405. PMLR, 2025. URL https://proceedings.mlr.press/v267/takezawa25a.html.

Michael Valancius, Maxwell Lennon, and Junier Oliva. Acquisition conditioned oracle for nongreedy active feature acquisition. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 48957–48975. PMLR, 2024. URL https://proceedings.mlr.press/v235/valancius24a.html.

Honghao Wei, Xiyue Peng, Arnob Ghosh, and Xin Liu. Adversarially trained weighted actor-critic for safe offline reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-1673.

Hongming Zhang, Tongzheng Ren, Chenjun Xiao, Dale Schuurmans, and Bo Dai. Provable representation with efficient planning for partially observable reinforcement learning. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 59759–59782. PMLR, 2024. URL https://proceedings.mlr.press/v235/zhang24bd.html.

Guide to the appendices. The model and information structure are in Appendix A. For the core theory, Appendix C proves the two-party example, Appendix L gives the common-response identity and constrained-geometry results, and Appendix F proves both certificate results. Appendices B and E provide the testing and representation bounds. Appendix D treats bounded intervention reach. For the main experiments, Appendix M specifies the controlled tasks, while Appendices G and N give the resource-allocation model, record diagnostic, and certificate evaluation. The additional protocol and query-scheduling studies are in Appendices H–K: intervention search, reference anchoring, timing and transfer, and response-protocol comparison, respectively.

A Formal Model and Information Structure

This appendix develops the model in Section 3. It specifies the information available at each decision, the resource and feasibility constraints, and the policy classes used to define the recoverability gap. The final subsection shows how the continuous convex construction and its sequential extension fit this model.

A.1 Public Protocol and Private Instance

Let Θ be the private instance space, $\mathcal{S}$ the public state space, and $\mathcal{Y}$ the response space. We use standard Borel spaces for these objects and for the ambient query and action spaces. This includes the finite and Euclidean spaces used in the paper and allows conditional distributions to be represented by measurable probability kernels. The public protocol is

$$I_{\mathrm{pub}} = \big(H, \mathcal{S}, \rho_0, \{\mathcal{Q}_t\}_{t=0}^{H-1}, \{\mathcal{A}_t\}_{t=0}^{H-1}, \mathcal{Y}, \{P_t\}_{t=0}^{H-1}, c, B, Q, K, \pi_{\mathrm{ref}}, \mathcal{F}, \nu\big). \tag{19}$$

Here H is a positive integer and ρ_0 is the initial public-state law. The query limit Q and intervention quota K are nonnegative integers. The admissible query and action sets, $\mathcal{Q}_t(\tau_t^-)$ and $\mathcal{A}_t(\tau_t^+)$, may depend on the respective transcripts. We assume that these set-valued maps are measurable. The kernels P_t describe public state transitions. The cost rule $c = (c_t)_{t=0}^{H-1}$, resource budget B, query limit Q, intervention quota K, and terminal feasible set $\mathcal{F} \subseteq \mathcal{S}$ apply to both policy classes.

The complete instance is $I = (I_{\mathrm{pub}}, \theta)$, where $\theta \sim \nu$ and ν is the public evaluation prior on Θ. There are M parties. Party m observes its local private information θ_m, which includes its local objective and is a measurable function of θ. The private instance determines the local objectives, response kernels $O_{\theta,t}$, and terminal loss ℓ_θ. The rules specifying admissible queries and actions, costs, and public state transitions are common to all instances. The reference π_{ref} is a deterministic measurable transcript policy. It specifies both a query rule and an action rule. $\pi_{\mathrm{ref},t}$ denotes its action rule at stage t. We assume that the reference satisfies the resource and terminal feasibility conditions below for every $\theta \in \Theta$ and every budget under consideration. These assumptions ensure that both admissible policy classes are nonempty. Table 2 summarizes the notation used in this appendix.

Table 2: Notation for the formal model.

Symbol	Meaning
$I = (I_{\mathrm{pub}}, \theta)$	Public protocol and private instance.
Θ, ν	Private instance space and public prior.
M, θ_m	Number of parties and party m's local private information.
H, S_t	Protocol horizon and public coordination state at stage t.
τ_t^-, τ_t^+	Transcripts before a query and after its response.
q_t, Y_t, A_t	Query, returned response, and coordination action.
$O_{\theta,t}, P_t$	Private-response kernel and public state-transition kernel.
Q, B, K	Query limit, resource budget, and intervention quota.
Q_t, B_t, K_t	Remaining query allowance, resources, and intervention quota.
$\pi_{\mathrm{ref}}, A_t^{\mathrm{ref}}$	Reference policy and its action at the current transcript.
Z_t^π	Indicator that A_t^π differs from A_t^{ref}.
$\mathcal{F}$	Public terminal feasible set.
U	Independent random tape implementing policy randomization.
$\Pi_{\mathrm{tr}}, \Pi_{\mathrm{full}}$	Transcript policies and instance-aware planners.
$L_\theta(\pi)$	Expected terminal loss of policy π on instance θ.

Table 2 continued.

Symbol	Meaning
$V_{\mathrm{tr}}^{(K)}, V_{\mathrm{full}}^{(K)}$	Optimal transcript-policy and instance-aware Bayes losses.
V_{ref}	Bayes loss of the reference policy.
$\mathrm{Gap}_{\mathrm{rec}}^{(K)}$	Recoverability gap between the two optimal Bayes losses.
$\mathrm{Adv}_{\mathrm{tr}}^{(K)}, \mathrm{Adv}_{\mathrm{full}}^{(K)}$	Improvements relative to the reference.
$\mathcal{R}_{\mathrm{mm}}^{(K)}$	Minimax transcript excess over the per-instance planner value.
$\theta_0 \equiv_{\Pi,H} \theta_1$	Transcript equivalence under the common policy class Π.
d_{TV}^{Π}	Largest transcript total variation over policies in Π.
$\mathcal{E}_{\mathcal{B}}^{(K)}(\mathcal{G})$	Smallest worst-case excess Bayes loss across budgets.
C_t, B_C	Additional decision certificate and its separate bit budget.

A.2 Interaction Sequence

The initial state has law ρ_0, independently of θ. Let B_t, Q_t, and K_t denote the remaining resource budget, query allowance, and intervention quota before the stage-t query. Initially,

$$\tau_0^- = (S_0, B_0, Q_0, K_0), \qquad S_0 \sim \rho_0, \quad (B_0, Q_0, K_0) = (B, Q, K). \tag{20}$$

At stage $t = 0, \dots, H-1$, the coordinator selects $q_t \in \mathcal{Q}_t(\tau_t^-)$ and receives

$$Y_t \sim O_{\theta,t}(\cdot \mid \tau_t^-, q_t), \qquad \tau_t^+ = (\tau_t^-, q_t, Y_t). \tag{21}$$

It then selects a coordination action $A_t \in \mathcal{A}_t(\tau_t^+)$, and the state evolves as $S_{t+1} \sim P_t(\cdot \mid S_t, q_t, Y_t, A_t)$. The state S_t contains the public variables required by this transition rule. A policy may use the full transcript even when the current state is a smaller summary. The remaining resource budget, query allowance, and intervention quota are updated according to the public rules in Appendix A.5. The next transcript is $\tau_{t+1}^- = (\tau_t^+, A_t, S_{t+1}, B_{t+1}, Q_{t+1}, K_{t+1})$. Thus the transcript retains all preceding states, queries, responses, actions, and resource updates. The public information available before the query and after the response is represented by the sigma-fields

$$\mathcal{F}_t^- = \sigma(\tau_t^-), \qquad \mathcal{F}_t^+ = \sigma(\tau_t^+). \tag{22}$$

The terminal transcript τ_H^- contains the final state S_H and every earlier transcript prefix. A stage used only to obtain a response can be represented by a singleton action set, so there is no choice of coordination action at that stage. If the protocol allows an unused query slot, it includes a null query $q^{\emptyset}$ with a fixed uninformative response. This permits a terminal action after fewer than Q queries. Early stopping can be represented by an absorbing state with null queries and no further interventions through the remaining stages. In the query-scheduling experiments of Section 6, A_t selects the party and the amount of local work for the next query q_{t+1}. The public state records this choice, which determines the query rule at stage $t+1$. A departure from the reference choice of the next query therefore counts as an intervention at stage t. The initialization queries form a fixed prefix whose responses are included in the transcript. Each query is charged when it is issued. For a fixed transcript policy, the response kernels are the only factors in the interaction law that depend on θ. Its query and action rules use the observed history, and its state transitions follow the common public kernels. This property yields the transcript-equivalence argument in Appendix A.7.

A.3 Available Information

Table 3 summarizes the information supplied during coordination and offline evaluation. Both policy classes observe the public history, and the instance-aware planner also observes θ. Party m observes the public history and its local private information θ_m.

Knowledge of θ allows the planner to compute quantities determined by the private instance, the public model, and the observed history. For example, it can compute the response to a given query when the response rule is deterministic. In stochastic protocols, both policy classes receive the realized responses and public states as the interaction proceeds. A policy learned offline belongs to Π_{tr} when its query and action rules depend on θ only through the public transcript. The offline evaluator in Section 6 uses private objectives to compute benchmarks and training labels. In Section 6.3, the evaluator also selects an alternative query at each decision state. The learned rule uses transcript features to decide whether to adopt this query. The reported capture measures the improvement achieved by this decision rule when the alternative queries are supplied by the evaluator.

Table 3: Information supplied during coordination and offline evaluation.

Information	Party m	Transcript policy	Instance-aware planner	Offline evaluator
Public protocol, state, and resource limits	✓	✓	✓	✓
Public interaction history	✓	✓	✓	✓
Private information θ_m	✓	–	✓	✓
Full private instance θ	–	–	✓	✓
Public query and action sets	✓	✓	✓	✓
Inputs for computation	Public inputs and θ_m	Public inputs	Public inputs and θ	Public inputs and θ
Random response and transition outcomes	As observed	As observed	As observed	Recorded or simulated
Evaluator-generated continuation labels	–	–	–	✓
Certificate in the augmented interface	If sender	After receipt	After receipt	✓

A check mark denotes supplied information. Policy entries refer to the evaluated interaction. Evaluator-generated labels may be used in offline training. Each role may compute functions of its available inputs.

A.4 Randomized Policy Classes

A policy consists of query kernels and action kernels over the finite horizon. We write π for the complete policy and π_t for its stage-t action kernel. A transcript policy uses

$$q_t \sim \mu_t(\cdot \mid \tau_t^-), \qquad A_t \sim \pi_t(\cdot \mid \tau_t^+). \tag{23}$$

The kernels are measurable and supported on the public query and action sets at the corresponding histories. We denote the class of all such policies by Π_{tr}. An instance-aware planner may additionally condition both kernels on the private instance:

$$q_t \sim \mu_t^{\mathrm{full}}(\cdot \mid \theta, \tau_t^-), \qquad A_t \sim \pi_t^{\mathrm{full}}(\cdot \mid \theta, \tau_t^+). \tag{24}$$

The resulting class is Π_{full}. Both policy classes use the same public query and action sets and interact through the same response and transition kernels. Equivalently, a policy can be represented by measurable functions of the available information and a random tape U. The tape consists of independent random variables, each uniform on $[0,1]$, and is independent of θ, S_0, and the randomness generating responses and public transitions. For a transcript policy, write $q_t = g_t(\tau_t^-, U)$ and $A_t = h_t(\tau_t^+, U)$, with a separate coordinate of U used for each query and action. The conditional distributions of q_t and A_t are the kernels in Eq. (23). For each θ and policy π, the initial-state law and the successive kernels define an interaction law $\mathbb{P}_\theta^\pi$. Expectations under this law are denoted by $\mathbb{E}_\theta^\pi$. The same transcript kernels are used for every θ. An instance-aware planner uses the additional conditioning in Eq. (24). Since the latter may ignore θ, $\Pi_{\mathrm{tr}} \subseteq \Pi_{\mathrm{full}}$.

A.5 Budgets and Sparse Interventions

All almost-sure resource and feasibility requirements in this appendix are imposed under $\mathbb{P}_\theta^\pi$ for every $\theta \in \Theta$. The prior ν determines how losses are averaged across these instances. For the nonnegative public stage cost c_t, the cumulative resource use is

$$C_H^\pi = \sum_{t=0}^{H-1} c_t(\tau_t^+, q_t, A_t). \tag{25}$$

A policy is budget feasible if $C_H^\pi \le B$ almost surely. The cost and budget may be vectors, in which case the inequality holds componentwise. This includes separate communication and local-work budgets. The stage cost may depend on the public history as well as the current query and action. When the protocol includes the null query $q^\emptyset$, the query count is $N_{H,\mathrm{qry}}^\pi = \sum_{t=0}^{H-1} \mathbf{1}\{q_t \neq q^\emptyset\} \le Q$. An unused query slot has zero query cost. Any separate cost of the coordination action is included in c_t. If the protocol has no null query, the query count is H, and we may take $Q = H$. We keep H and Q fixed in each comparison and suppress them in the policy-class and value notation, as in Section 3.3. At the current post-response transcript, define

$$A_t^{\mathrm{ref}} = \pi_{\mathrm{ref},t}(\tau_t^+), \qquad Z_t^\pi = \mathbf{1}\{A_t^\pi \neq A_t^{\mathrm{ref}}\}. \tag{26}$$

Here A_t^π denotes the action chosen by policy π, and an intervention occurs when $Z_t^\pi = 1$. At every stage, the reference rule is evaluated at the transcript generated by π, including any earlier interventions and resource use. The total number of interventions is $N_{H,\mathrm{int}}^\pi = \sum_{t=0}^{H-1} Z_t^\pi$. A policy is K-sparse if $N_{H,\mathrm{int}}^\pi \le K$ almost surely. The query limit Q bounds the number of issued non-null queries, and K bounds the number of departures from the reference action rule. In the query-scheduling representation, an intervention occurs when A_t selects a next query different from the reference choice. The query counts toward Q when it is issued. We write $\Pi_{\mathrm{tr}}^{(K)}(B)$ and $\Pi_{\mathrm{full}}^{(K)}(B)$ for the policies in the respective information classes that satisfy the resource budget, query limit, and intervention quota. Before the stage-t query, B_t is the initial budget minus the accumulated cost, Q_t is Q minus the number of non-null queries issued, and K_t is K minus the number of interventions. These quantities are determined by the public history.

A.6 Loss, Feasibility, and Values

The terminal feasibility requirement is

$$S_H^\pi \in \mathcal{F} \qquad \mathbb{P}_\theta^\pi\text{-almost surely for every } \theta \in \Theta. \tag{27}$$

A projection or other correction used to enforce terminal feasibility is included in the public transition rule shared by both policy classes. The terminal loss is evaluated at the resulting state. In Section 6, the raw balance violation is a separate diagnostic measured before the final balance projection. Let $\Pi_{\mathrm{tr,feas}}^{(K)}(B)$ and $\Pi_{\mathrm{full,feas}}^{(K)}(B)$ denote the subclasses of $\Pi_{\mathrm{tr}}^{(K)}(B)$ and $\Pi_{\mathrm{full}}^{(K)}(B)$, respectively, that also satisfy Eq. (27). The loss $(\theta, s) \mapsto \ell_\theta(s)$ is jointly measurable. We assume that it is bounded below by a finite measurable function $\underline{\ell}(\theta)$ with $\mathbb{E}_{\theta\sim\nu}|\underline{\ell}(\theta)| < \infty$. The reference has finite expected loss on every instance and finite Bayes loss. These assumptions ensure that the two optimal Bayes losses are finite. For policy π, its expected terminal loss on instance θ is

$$L_\theta(\pi) = \mathbb{E}_\theta^\pi[\ell_\theta(S_H)]. \tag{28}$$

The Bayes values are those in Eq. (1).

$$V_{\mathrm{full}}^{(K)}(\nu, B) = \inf_{\pi\in\Pi_{\mathrm{full,feas}}^{(K)}(B)} \mathbb{E}_{\theta\sim\nu} L_\theta(\pi), \qquad V_{\mathrm{tr}}^{(K)}(\nu, B) = \inf_{\pi\in\Pi_{\mathrm{tr,feas}}^{(K)}(B)} \mathbb{E}_{\theta\sim\nu} L_\theta(\pi). \tag{29}$$

For the fixed prior ν, the infimum ranges over transcript policies whose query and action kernels are shared across instances. Each such policy adapts through the realized transcript. The optimal instance-aware value is the oracle benchmark. Because the constrained transcript class is contained in the constrained instance-aware class, and the reference belongs to both,

$$V_{\mathrm{full}}^{(K)}(\nu, B) \le V_{\mathrm{tr}}^{(K)}(\nu, B) \le V_{\mathrm{ref}}(\nu, B), \qquad V_{\mathrm{ref}}(\nu, B) = \mathbb{E}_{\theta\sim\nu} L_\theta(\pi_{\mathrm{ref}}). \tag{30}$$

The recoverability gap in Eq. (2) is therefore nonnegative. The improvements over the reference policy are

$$\mathrm{Adv}_{\mathrm{full}}^{(K)}(\nu, B) = V_{\mathrm{ref}}(\nu, B) - V_{\mathrm{full}}^{(K)}(\nu, B), \quad \mathrm{Adv}_{\mathrm{tr}}^{(K)}(\nu, B) = V_{\mathrm{ref}}(\nu, B) - V_{\mathrm{tr}}^{(K)}(\nu, B). \tag{31}$$

Their difference gives Eq. (3). This decomposition expresses the recoverability gap as the part of the optimal instance-aware improvement that transcript policies cannot recover. For the minimax comparison, fix a nonempty instance family $\Theta_0 \subseteq \Theta$. Define the per-instance planner value

$$L_\theta^{\mathrm{full},\star} = \inf_{\pi\in\Pi_{\mathrm{full,feas}}^{(K)}(B)} L_\theta(\pi), \tag{32}$$

where the fixed B and K are suppressed on the left. The minimax excess loss of transcript policies is

$$\mathcal{R}_{\mathrm{mm}}^{(K)}(\Theta_0, B) = \inf_{\pi\in\Pi_{\mathrm{tr,feas}}^{(K)}(B)} \sup_{\theta\in\Theta_0} \big[L_\theta(\pi) - L_\theta^{\mathrm{full},\star}\big]. \tag{33}$$

Here a single transcript policy is compared with the planner value on each instance. This is the prior-free comparison used alongside the Bayes gap in Theorems 4.1 and D.1.

A.7 Exact and Approximate Transcript Aliasing

Fix two private instances θ_0, θ_1 and a nonempty class $\Pi \subseteq \Pi_{\mathrm{tr}}$ whose policies satisfy the resource, query, intervention, and terminal feasibility constraints under both instances. For a fixed $\pi \in \Pi$, let $P^{\pi,-}_{\theta,t}$ and $P^{\pi,+}_{\theta,t}$ denote the laws of τ_t^- and τ_t^+ under $\mathbb{P}^\pi_\theta$. We write $\theta_0 \equiv_{\Pi,H} \theta_1$ when, for every $\pi \in \Pi$,

$$P^{\pi,-}_{\theta_0,t} = P^{\pi,-}_{\theta_1,t}, \quad t = 0, \dots, H, \qquad P^{\pi,+}_{\theta_0,t} = P^{\pi,+}_{\theta_1,t}, \quad t = 0, \dots, H-1. \tag{34}$$

The same policy is used on both instances. Section 3.4 uses Π_{tr} for the class of all transcript policies admissible under both instances for the fixed public protocol. The definition here also allows a nonempty subclass Π. A sufficient condition for equivalence is that, for every $\pi \in \Pi$ and stage t, the two response kernels agree at (τ_t^-, q_t) almost surely under both $\mathbb{P}^\pi_{\theta_0}$ and $\mathbb{P}^\pi_{\theta_1}$. Fix $\pi \in \Pi$. The initial transcript has the same law under both instances. If the pre-query transcripts have the same law, applying the common query kernel gives the same law for (τ_t^-, q_t). Equality of the response kernels then gives the same post-response law. The common action and transition kernels, together with the public resource updates, give the same law for τ_{t+1}^-. Induction proves Eq. (34). For probability laws P and P' on the same measurable space, write $\mathrm{TV}(P, P') = \sup_E |P(E) - P'(E)|$. Since τ_H^- contains the whole interaction, define

$$d^{\Pi}_{\mathrm{TV}}(\theta_0, \theta_1) = \sup_{\pi \in \Pi} \mathrm{TV}\big(P^{\pi,-}_{\theta_0,H}, P^{\pi,-}_{\theta_1,H}\big). \tag{35}$$

For $\varepsilon \in [0,1]$, the pair is ε-aliased under Π if $d^{\Pi}_{\mathrm{TV}}(\theta_0, \theta_1) \le \varepsilon$. Exact equivalence is the case $\varepsilon = 0$. Each prefix is a measurable function of the terminal transcript, so total variation at a prefix is at most the total variation of the complete transcript. More generally, fix $\pi \in \Pi$ and let W be a decision record obtained from the available transcript by the same measurable map under both instances.

$$\mathrm{TV}\big(P^{\pi,W}_{\theta_0}, P^{\pi,W}_{\theta_1}\big) \le \mathrm{TV}\big(P^{\pi,-}_{\theta_0,H}, P^{\pi,-}_{\theta_1,H}\big) \le d^{\Pi}_{\mathrm{TV}}(\theta_0, \theta_1), \tag{36}$$

where $P^{\pi,W}_\theta$ denotes the law of W. For an ε-aliased pair, the pre-action record laws therefore have total variation at most ε for every $\pi \in \Pi$. This gives the bound on record distinguishability used in Corollaries B.3 and C.4.

A.8 Uniform Cross-Budget Recoverability

Let $\mathcal{B}$ be a nonempty measurable set of public resource budgets. We keep the instance family, prior ν, horizon H, query limit Q, and intervention quota K fixed while varying $B \in \mathcal{B}$. The public protocol is specified by common rules across budgets. For example, the set of affordable queries and the reference continuation may depend on the observed budget. A budget-aware policy uses query and action rules that are jointly measurable in the transcript and the budget. We write $\pi(\tau, B)$ for these rules and π_B for the policy obtained by fixing B. Let $\mathcal{G}$ be a nonempty family of such policies satisfying $\pi_B \in \Pi^{(K)}_{\mathrm{tr,feas}}(B)$ for every $\pi \in \mathcal{G}$, $B \in \mathcal{B}$. Let $L_{\theta,B}(\pi)$ denote the expected terminal loss of π_B on instance θ at budget B. Define

$$\mathcal{E}^{(K)}_{\mathcal{B}}(\mathcal{G}) = \inf_{\pi \in \mathcal{G}} \sup_{B \in \mathcal{B}} \Big[\mathbb{E}_{\theta \sim \nu} L_{\theta,B}(\pi) - V^{(K)}_{\mathrm{full}}(\nu, B)\Big]. \tag{37}$$

For $\epsilon \ge 0$, the instance-aware benchmark is ϵ-uniformly recoverable over $\mathcal{B}$ within $\mathcal{G}$ if this quantity is at most ϵ. The supremum measures a policy's largest excess Bayes loss across budgets. The infimum ranges over the budget-aware policies in $\mathcal{G}$, each defined on all of $\mathcal{B}$. For finitely many budgets, an unrestricted measurable rule can assign a separately chosen policy to each budget. Restrictions on the representation or function class may limit which policies can be combined in this way. Appendix E studies action rules that use the representation $\phi(\tau, B)$. Eq. (37) describes the optimal performance within $\mathcal{G}$ across the budget set. Section 6.3 examines how predictors trained at some budgets perform at held-out budgets.

A.9 Certificate-Augmented Policies

A decision certificate is an additional locally computed message received after an ordinary response and before a coordination action. We first specify the information available when certificates can be sent at several stages, then specialize to the single-decision setting of Section 5. Let C_t

be the combined certificate received at stage t. Before the stage-t query, the coordinator retains $\widetilde{\tau}_t^- = (\tau_t^-, C_0, \dots, C_{t-1})$. After the ordinary response, the record available to a certificate sender is

$$W_t = (\tau_t^+, C_0, \dots, C_{t-1}). \tag{38}$$

Thus earlier certificates can inform later queries, messages, and actions. The protocol specifies the senders, the stages at which they may send, and the message alphabets and encodings. Each designated party uses a measurable kernel

$$C_{m,t} \sim K_{m,t}(\cdot \mid \theta_m, W_t). \tag{39}$$

Each sender samples its message using fresh local randomness, independent of the policy and environment random sources. At a stage with several senders, their draws are conditionally independent given θ and W_t. The combined message C_t lists the messages in the public sender order. The record available for the coordination action is $\widetilde{\tau}_t^+ = (W_t, C_t)$. The query and action rules use $\widetilde{\tau}_t^-$ and $\widetilde{\tau}_t^+$, respectively:

$$q_t \sim \widetilde{\mu}_t(\cdot \mid \widetilde{\tau}_t^-), \qquad A_t \sim \widetilde{\pi}_t(\cdot \mid \widetilde{\tau}_t^+). \tag{40}$$

The ordinary response law remains $O_{\theta,t}(\cdot \mid \tau_t^-, q_t)$, and the query sets, action sets, public transitions, and ordinary cost rule are those of I_{pub}. In the augmented interaction, $\tau_t^{\pm}$ denotes the ordinary record obtained by omitting the certificate messages. The reference action remains $\pi_{\mathrm{ref},t}(\tau_t^+)$, evaluated at that record, so interventions are counted as in Eq. (26). Let $b_{C,t}(C_t)$ denote the total encoded length of C_t in bits. The certificate budget is

$$\sum_{t=0}^{H-1} b_{C,t}(C_t) \le B_C \qquad \mathbb{P}_\theta^\pi\text{-almost surely for every } \theta \in \Theta. \tag{41}$$

The budget B_C is separate from B. The public record W_t determines which parties send at stage t. If no party sends, C_t is a fixed empty message and $b_{C,t}(C_t) = 0$. Each transmitted message uses its specified fixed-length or self-delimiting binary code. A self-delimiting code allows the receiver to identify the end of a message from the encoded bits. An augmented policy consists of the public transmission rules, the local certificate kernels, and the coordinator's query and action kernels. Let $\Pi^{(K)}_{\mathrm{tr}+C,\mathrm{feas}}(B, B_C)$ be the class of these policies satisfying both budgets, the query limit, the intervention quota, and terminal feasibility. Its optimal Bayes loss is

$$V^{(K)}_{\mathrm{tr}+C}(\nu; B, B_C) = \inf_{\pi \in \Pi^{(K)}_{\mathrm{tr}+C,\mathrm{feas}}(B,B_C)} \mathbb{E}_{\theta\sim\nu} L_\theta(\pi). \tag{42}$$

This infimum optimizes the allowed local message rules together with the coordinator's decisions. A fixed certificate scheme Γ specifies its public transmission rules and local message kernels. Write $V^{(K)}_{\mathrm{tr}+C}(\nu; B, B_C \mid \Gamma)$ for the same infimum restricted to these message rules. We call Γ ϵ-sufficient under ν if

$$V^{(K)}_{\mathrm{tr}+C}(\nu; B, B_C \mid \Gamma) - V^{(K)}_{\mathrm{full}}(\nu, B) \le \epsilon. \tag{43}$$

The one-bit scheme in Section 5.1 attains the instance-aware optimum on each of the two instances. Every ordinary transcript policy can be used in the augmented class by sending no certificates. An instance-aware planner can simulate the local message kernels from θ and independent randomness and then follow the same query and action rules. Consequently,

$$V^{(K)}_{\mathrm{full}}(\nu, B) \le V^{(K)}_{\mathrm{tr}+C}(\nu; B, B_C) \le V^{(K)}_{\mathrm{tr}}(\nu, B). \tag{44}$$

The remaining recoverability gap is $V^{(K)}_{\mathrm{tr}+C}(\nu; B, B_C) - V^{(K)}_{\mathrm{full}}(\nu, B)$. The reduction relative to the ordinary transcript class is $V^{(K)}_{\mathrm{tr}}(\nu, B) - V^{(K)}_{\mathrm{tr}+C}(\nu; B, B_C)$. Section 5 considers a single certificate with intervention quota $K = 1$. At the certificate stage, $W = \tau_t^+$ and $C = C_t$, since no earlier certificates have been sent. In the construction of Section 5.1, party 1 computes the bit in Eq. (17) from its local objective and the public curvature values. It sends this bit immediately before the terminal allocation, using $B_C = 1$. For this noiseless construction, we identify C_t with $C_{1,t}$, the bit denoted by C_1 in the main text. Theorem 5.1 gives the resulting optimal allocation. In Corollary F.4, C_t is the received bit $\widetilde{C}_1$, obtained after an independent flip with probability p.

A.10 The Convex Construction as a Coordination Protocol

Theorem 4.1 uses the objectives in Eq. (4), the swapped assignments in Eq. (5), and the feasible set $\mathcal{X}$ in Eq. (6). The two instances agree on $[0, 1/2]$ for each local objective. Permitted price queries return optimizers in $[0, 1/2]$, while the terminal constraint requires both allocation coordinates to lie in $[1/2, 1]$. For a query limit $Q \geq 0$, take $H = Q + 1$. The first Q stages are query slots, and stage Q chooses the terminal allocation. At an ordinary query slot, the coordinator either issues a party-specific price query $q_t = (m, \lambda)$ or uses $q^{\emptyset}$. A price query returns $Y_t = \mathcal{R}_{m,\theta}(\lambda)$, $m \in \{1, 2\}$, $\lambda \in [-1/4, 1/4]$, with the response defined in Eq. (7). At each query stage, there is a single coordination action, which leaves the allocation unchanged. At stage Q, the query is null and the action is an allocation $A_Q \in \mathcal{X}$. Null queries return a fixed empty response. This also covers $Q = 0$, when the sole stage chooses the allocation. Use $\mathcal{S} = \mathcal{X}$ and initialize $S_0 = x^{\mathrm{ref}} = (3/4, 3/4)$. The public transition preserves S_t during the query slots and sets $S_{Q+1} = A_Q$ at the terminal stage. The feasible set is $\mathcal{F} = \mathcal{X}$, and the terminal loss is $\ell_\theta(S_{Q+1}) = f_{1,\theta}(S_{Q+1,1}) + f_{2,\theta}(S_{Q+1,2})$.

Table 4: Specialization of the formal model to Theorem 4.1.

Model component	Specialization
Private instance and prior	Assignment of κ_L and κ_H to the two parties; uniform prior on the two assignments.
Horizon and initial state	$H = Q + 1$ and $S_0 = (3/4, 3/4)$.
Query slots	At most Q prices, each with a chosen party and $\lambda \in [-1/4, 1/4]$.
Response	Exact local optimizer $\mathcal{R}_{m,\theta}(\lambda) = 1/4 - \lambda$.
Coordination actions	A single action leaving the allocation unchanged during query slots; a terminal allocation $A_Q \in \mathcal{X}$.
Public transition	Preserve the state during querying and set $S_{Q+1} = A_Q$ at the terminal stage.
Ordinary resources	One unit per price query, $B = Q$, and 0 resource cost for allocation actions.
Reference	Null queries and terminal allocation $x^{\mathrm{ref}} = (3/4, 3/4)$.
Intervention quota	$K = 1$; an intervention occurs if $A_Q \neq x^{\mathrm{ref}}$.
Terminal feasibility	$\mathcal{F} = \mathcal{X} = \{x \in [0, 1]^2 : x_1 + x_2 = 3/2\}$.

Assign unit resource cost to each price query and zero resource cost to null queries and allocation actions. Setting $B = Q$ then makes the resource constraint coincide with the query-count constraint. Table 4 summarizes this specialization. By Eq. (8), a fixed transcript policy receives the same response to each permitted query under the two instances. Its terminal allocation therefore has the same law under both instances. The instance-aware planner knows the curvature assignment and can choose the corresponding optimum in Eq. (9). Under the uniform prior on the two instances, the symmetric reference attains the optimal transcript value, giving the recoverability gap in Eq. (10). Recall that $\Sigma = \kappa_L + \kappa_H$ and $d = (\kappa_H - \kappa_L)/(4\Sigma)$. The sequential extension in Appendix D uses a horizon $H \geq K$ and an increment bound $0 < \delta \leq d$. Take $\mathcal{S} = \mathcal{F} = [-1/4, 1/4]$ and $S_0 = z_0 = 0$. At each stage, the coordinator issues a price query or a null query, receives the response, and then chooses the increment $A_t = u_t$. The public state is $S_t = z_t$, and admissible increments satisfy $|u_t| \leq \delta$ and $z_t + u_t \in [-1/4, 1/4]$. The public transition is $z_{t+1} = z_t + u_t$. The terminal loss is

$$\ell_\theta(z_H) = f_{1,\theta}(3/4 + z_H) + f_{2,\theta}(3/4 - z_H). \tag{45}$$

The reference action is $u_t = 0$, so each nonzero increment counts as one intervention. The resource budget B and query limit Q apply only to price queries. Increments have zero ordinary resource cost. The response rule in Eq. (7) does not depend on z_t, so Eq. (8) holds throughout the sequence. Earlier increments determine the current displacement and the remaining intervention quota. The term $-\Sigma z u$ in Eq. (98) shows how the current displacement changes the gain from a later increment u. This gives the protocol of Theorem D.1.

B General Aliasing Bounds

This appendix gives the classical two-point testing argument (Le Cam, 1986; Chen et al., 2024), its unequal-prior and conditional versions, and its application to representations of a transcript.

Consider a decision problem with action set $\mathcal{A} = \{a_0, a_1\}$ and two equiprobable private instances θ_0, θ_1. A policy observes a record W, whose law under θ_i is P_i. Let $J_i(a)$ denote the loss of action a under θ_i. Suppose

$$J_0(a_1) - J_0(a_0) \geq \Delta, \qquad J_1(a_0) - J_1(a_1) \geq \Delta \tag{46}$$

for some $\Delta > 0$. An instance-aware planner chooses a_i under θ_i. Relative to this benchmark, a policy π selecting $A \in \mathcal{A}$ has Bayes regret $\mathrm{Reg}(\pi) = \frac{1}{2}\sum_{i=0}^{1}\big(\mathbb{E}^{\pi}_{\theta_i}[J_i(A)] - J_i(a_i)\big)$, where the expectation is over W and the policy's randomization.

Theorem B.1 (Two-action transcript aliasing)**.** *Let θ_0, θ_1 be two equiprobable private instances. The observed record W has law P_i under θ_i. A randomized decision rule π based on W selects an action $A \in \mathcal{A} = \{a_0, a_1\}$. Let $J_i(a)$ denote the loss of action a under θ_i. Suppose the margins in Eq.* (46) *hold for some $\Delta > 0$. Every such rule has Bayes regret*

$$\mathrm{Reg}(\pi) = \frac{1}{2}\sum_{i=0}^{1}\big(\mathbb{E}^{\pi}_{\theta_i}[J_i(A)] - J_i(a_i)\big) \geq \frac{\Delta}{2}\big(1 - \mathrm{TV}(P_0, P_1)\big). \tag{47}$$

Proof. Write $p(w) = \pi(a_1 \mid w)$ for the probability of choosing a_1 after observing w. The same measurable function p, with values in $[0, 1]$, describes the decision rule under both instances. Its two error probabilities are

$$\mathbb{P}^{\pi}_{\theta_0}(A = a_1) = \int p\, dP_0, \qquad \mathbb{P}^{\pi}_{\theta_1}(A = a_0) = \int (1 - p)\, dP_1. \tag{48}$$

For any measurable function g with $0 \leq g \leq 1$, the definition of total variation gives

$$\left|\int g\, dP_0 - \int g\, dP_1\right| \leq \mathrm{TV}(P_0, P_1). \tag{49}$$

Indeed, write $g(w) = \int_0^1 \mathbf{1}\{g(w) > u\}\, du$ and apply the definition of total variation to the events $\{g > u\}$ inside the integral. Taking $g = p$ yields

$$\mathbb{P}^{\pi}_{\theta_0}(A = a_1) + \mathbb{P}^{\pi}_{\theta_1}(A = a_0) = 1 + \int p\, dP_0 - \int p\, dP_1 \geq 1 - \mathrm{TV}(P_0, P_1). \tag{50}$$

Under θ_i, choosing a_i has zero regret and choosing the other action incurs regret at least Δ. Averaging the two instance-wise regrets proves Eq. (47). □

Corollary B.2 (Unequal priors and loss margins)**.** *In the same two-action problem, let $\Pr(\theta = \theta_0) = \rho \in [0, 1]$ and suppose*

$$J_0(a_1) - J_0(a_0) \geq \Delta_0, \qquad J_1(a_0) - J_1(a_1) \geq \Delta_1, \tag{51}$$

where $\Delta_0, \Delta_1 > 0$. Define the Bayes regret for this prior by

$$\mathrm{Reg}_{\rho}(\pi) = \rho\left(\mathbb{E}^{\pi}_{\theta_0}[J_0(A)] - J_0(a_0)\right) + (1 - \rho)\left(\mathbb{E}^{\pi}_{\theta_1}[J_1(A)] - J_1(a_1)\right). \tag{52}$$

Every randomized decision rule based on W satisfies

$$\mathrm{Reg}_{\rho}(\pi) \geq \min\{\rho\Delta_0, (1 - \rho)\Delta_1\}\big(1 - \mathrm{TV}(P_0, P_1)\big). \tag{53}$$

Proof. With $p(w) = \pi(a_1 \mid w)$ as above, the regret is at least

$$\rho\Delta_0 \int p\, dP_0 + (1 - \rho)\Delta_1 \int (1 - p)\, dP_1. \tag{54}$$

Since both integrals are nonnegative, this is at least

$$\min\{\rho\Delta_0, (1 - \rho)\Delta_1\}\left(\int p\, dP_0 + \int (1 - p)\, dP_1\right). \tag{55}$$

Applying Eq. (50) proves the bound. □

Corollary B.3 (Approximate aliasing over a policy class)**.** *Consider the two-action problem of Theorem B.1. Let $\Pi \subseteq \Pi_{\mathrm{tr}}$ be a nonempty class of policies satisfying the resource, query, intervention, and terminal feasibility constraints under both instances. For each $\pi \in \Pi$, let $P_{\theta_i}^{\pi,W}$ be the law of its record W before the action under θ_i. If, for some $\varepsilon \in [0,1]$,*

$$\sup_{\pi\in\Pi} \mathrm{TV}\big(P_{\theta_0}^{\pi,W}, P_{\theta_1}^{\pi,W}\big) \le \varepsilon, \tag{56}$$

then

$$\inf_{\pi\in\Pi} \mathrm{Reg}(\pi) \ge \frac{\Delta}{2}(1-\varepsilon). \tag{57}$$

Proof. Fix $\pi \in \Pi$. Its query and action rules before the decision induce the two record laws $P_{\theta_i}^{\pi,W}$. Applying Theorem B.1 to these laws gives regret at least $\Delta(1-\varepsilon)/2$. Since this lower bound holds for every $\pi \in \Pi$, taking the infimum of $\mathrm{Reg}(\pi)$ over Π proves the claim. □

Appendix A.7 defines approximate aliasing using the complete transcript. Since W is obtained from the available transcript, Eq. (36) gives the required bound on its record laws. Under exact aliasing, we may take $\varepsilon = 0$, giving Bayes regret at least $\Delta/2$.

B.1 A Conditional Bound

In a sequential problem, the expected loss of an action may depend on the record observed before it is chosen. Fix the same transcript policy under both instances up to a given decision. Let W be the complete public record available then, with law P_i under θ_i. Compare two actions a_0, a_1 at this decision. After either action, follow a fixed transcript policy $\bar{\pi}$ for the remaining interaction. This policy may adapt to subsequent observations. Each action followed by $\bar{\pi}$ must satisfy the remaining resource, query, and intervention limits and the terminal feasibility constraint under both instances.

Let $J_i(w,a)$ be the expected terminal loss under θ_i when the record is w, action a is taken, and $\bar{\pi}$ is followed thereafter. This expectation includes future responses and state transitions. Assume that $J_i(\cdot,a)$ is measurable and integrable under P_i for each $i \in \{0,1\}$ and $a \in \{a_0,a_1\}$. Let $\mathcal{E}$ be a measurable set of decision records with $P_0(\mathcal{E}) > 0$ and $P_1(\mathcal{E}) > 0$. Conditioning on the public event $W \in \mathcal{E}$ gives the record laws

$$P_i^{\mathcal{E}}(D) = \frac{P_i(D \cap \mathcal{E})}{P_i(\mathcal{E})}, \qquad i \in \{0,1\}, \tag{58}$$

for measurable sets D of decision records. If $\Pr(\theta = \theta_0) = \rho \in (0,1)$, the conditional prior is

$$\rho_{\mathcal{E}} = \Pr(\theta = \theta_0 \mid W \in \mathcal{E}) = \frac{\rho P_0(\mathcal{E})}{\rho P_0(\mathcal{E}) + (1-\rho)P_1(\mathcal{E})}. \tag{59}$$

Corollary B.4 (Conditional transcript aliasing)**.** *Suppose that, for each $i \in \{0,1\}$, there is a constant $\Delta_{i,\mathcal{E}} > 0$ such that*

$$J_i(w, a_{1-i}) - J_i(w, a_i) \ge \Delta_{i,\mathcal{E}} \quad \text{for } P_i^{\mathcal{E}}\text{-almost every } w. \tag{60}$$

Let π use the specified rules before the decision and the continuation $\bar{\pi}$ afterward. Its current action $A \in \{a_0,a_1\}$ is selected by a randomized rule based on W. Define its conditional Bayes regret by

$$\begin{aligned}\mathrm{Reg}_{\mathcal{E}}(\pi) = \rho_{\mathcal{E}}\, \mathbb{E}_{\theta_0}^{\pi}\big[J_0(W,A) - J_0(W,a_0) \mid W \in \mathcal{E}\big] \\ + (1-\rho_{\mathcal{E}})\, \mathbb{E}_{\theta_1}^{\pi}\big[J_1(W,A) - J_1(W,a_1) \mid W \in \mathcal{E}\big].\end{aligned} \tag{61}$$

Its benchmark chooses a_i under θ_i and then follows the same continuation. Every randomized decision rule based on W satisfies

$$\mathrm{Reg}_{\mathcal{E}}(\pi) \ge \min\{\rho_{\mathcal{E}}\Delta_{0,\mathcal{E}}, (1-\rho_{\mathcal{E}})\Delta_{1,\mathcal{E}}\}\big(1 - \mathrm{TV}(P_0^{\mathcal{E}}, P_1^{\mathcal{E}})\big). \tag{62}$$

Proof. Let $p(w)$ be the probability that π chooses a_1 after observing w. Since the conditioning event is determined by W, the same function p describes the action rule under each conditional law. The pointwise margin conditions give

$$\mathrm{Reg}_{\mathcal{E}}(\pi) \ge \rho_{\mathcal{E}}\Delta_{0,\mathcal{E}} \int p\, dP_0^{\mathcal{E}} + (1-\rho_{\mathcal{E}})\Delta_{1,\mathcal{E}} \int (1-p)\, dP_1^{\mathcal{E}}. \tag{63}$$

Both integrals are nonnegative. Using the smaller coefficient as a common lower bound and applying Eq. (50) to $P_0^{\mathcal{E}}, P_1^{\mathcal{E}}$ proves the claim. □

The continuation losses may vary with w, while Eq. (60) requires the stated margins to hold almost surely under the corresponding conditional laws. If a history w has positive probability under both instances, taking $\mathcal{E} = \{w\}$ gives a bound at that history. For continuously distributed records, one can instead condition on a measurable set $\mathcal{E}$ with positive probability under both instances.

B.2 Data Processing

Return to the equiprobable two-action problem of Theorem B.1. Suppose the policy uses only a measurable representation $\phi(W)$ of its record. Let P_i^{ϕ} denote the law of $\phi(W)$ under θ_i. For every measurable set D in the representation space, $P_i^{\phi}(D) = P_i\big(\phi^{-1}(D)\big)$. Here $\phi^{-1}(D) = \{w : \phi(w) \in D\}$. Therefore,

$$\mathrm{TV}(P_0^{\phi}, P_1^{\phi}) = \sup_D \left|P_0\big(\phi^{-1}(D)\big) - P_1\big(\phi^{-1}(D)\big)\right| \le \mathrm{TV}(P_0, P_1). \tag{64}$$

Applying Theorem B.1 to the record $\phi(W)$ gives

$$\mathrm{Reg}(\pi) \ge \frac{\Delta}{2}\big(1 - \mathrm{TV}(P_0^{\phi}, P_1^{\phi})\big) \tag{65}$$

for every randomized decision rule based on that representation. Taking $W = (\tau, B)$ and $\Delta = \gamma$ gives the bound in Proposition E.1.

C The Continuous Convex Hard Pair

This appendix proves Theorem 4.1 and Corollary C.4. We derive the local response law and minimize the expected terminal loss over all randomized allocations in $\mathcal{X}$. We then examine how extensions of the response interface affect transcript aliasing. We use the objectives in Eq. (4), the two curvature assignments in Eq. (5), the feasible set in Eq. (6), and the response rule in Eq. (7). Recall that $0 < \kappa_L < \kappa_H$, $\Sigma = \kappa_L + \kappa_H$, $d = \frac{\kappa_H - \kappa_L}{4\Sigma} \in (0, 1/4)$. The intervention quota is $K = 1$. We suppress the fixed uniform prior and resource budget in the value notation, as in Section 4.1.

Theorem C.1 (Continuous convex recoverability gap, restated)**.** *For every finite query limit Q, each admissible adaptive randomized transcript policy induces the same law for every public transcript prefix under θ_0 and θ_1. The unique instance-aware optima are the allocations in Eq.* (9)*. Under the uniform prior,*

$$V_{\mathrm{tr}}^{(1)} = V_{\mathrm{ref}}, \qquad \mathrm{Gap}_{\mathrm{rec}}^{(1)} = V_{\mathrm{tr}}^{(1)} - V_{\mathrm{full}}^{(1)} = G = \frac{(\kappa_H - \kappa_L)^2}{32\Sigma}. \tag{66}$$

The minimax excess loss of transcript policies relative to the instance-aware optima also equals G.

C.1 Regularity and Local Responses

For $\kappa > 0$, differentiating the objective gives

$$\varphi'_\kappa(x) = \begin{cases} x - \dfrac{1}{4}, & 0 \le x \le \dfrac{1}{2}, \\ \dfrac{1}{4} + \kappa\left(x - \dfrac{1}{2}\right), & \dfrac{1}{2} < x \le 1. \end{cases} \tag{67}$$

The two branches of φ_κ have the same value $1/32$ and derivative $1/4$ at $x = 1/2$, so $\varphi_\kappa \in C^1([0,1])$. For $0 \le x < y \le 1$,

$$\min\{1, \kappa\}(y - x) \le \varphi'_\kappa(y) - \varphi'_\kappa(x) \le \max\{1, \kappa\}(y - x). \tag{68}$$

On the two branches, φ'_κ has slopes 1 and κ, respectively. When $x < 1/2 < y$, the bounds in Eq. (68) follow by adding the derivative increments on the two sides of $1/2$. Thus the derivative is Lipschitz with constant $\max\{1, \kappa\}$, and the objective is $\min\{1, \kappa\}$-strongly convex.

Lemma C.2 (Exact response identity)**.** *For every $\lambda \in [-1/4, 1/4]$ and every $\kappa > 0$,*

$$\arg\min_{x\in[0,1]} \{\varphi_\kappa(x) + \lambda x\} = \frac{1}{4} - \lambda. \tag{69}$$

Proof. The point $1/4 - \lambda$ lies in $[0, 1/2]$, where all objectives in the family agree. The derivative of $\varphi_\kappa(x) + \lambda x$ there is $x - 1/4 + \lambda$, which vanishes at this point. Adding a linear term preserves strong convexity, so the point is the unique minimizer on $[0, 1]$. □

Lemma C.3 (Transcript equivalence)**.** *For every finite Q, each admissible transcript policy for this construction induces identical laws for every public transcript prefix under θ_0 and θ_1. Its terminal allocation therefore has the same law under both instances.*

Proof. Fix a transcript policy π. Both instances start from the same public state x^{ref}. Suppose the pre-query transcript laws agree at a query slot. Applying the common query kernel gives the same joint law of the transcript and the selected query under the two instances. By Lemma C.2, the response to a party-specific price query depends only on its price. A null query returns the same empty response under both instances. The post-response transcript laws therefore agree. The common action and public transition rules, including the resource updates, preserve this equality at the next stage. Induction over the Q query slots gives identical laws for the record available at the terminal decision. The terminal action kernel is also common, so the allocation laws agree. This proves the prefix equivalence defined in Appendix A.7. □

C.2 Coordinated Optima and Exact Values

Proof of Theorem C.1. Every $x \in \mathcal{X}$ has both coordinates in $[1/2, 1]$. Set $y_i = x_i - 1/2$ for $i \in \{1, 2\}$, so $y_1 + y_2 = 1/2$ and $0 \le y_i \le 1/2$. For an assignment of curvatures (κ_1, κ_2) to the two parties, the terminal loss is

$$\ell_\theta(x) = \frac{1}{16} + \frac{1}{4}(y_1 + y_2) + \frac{\kappa_1}{2}y_1^2 + \frac{\kappa_2}{2}y_2^2 = \frac{3}{16} + \frac{\kappa_1}{2}y_1^2 + \frac{\kappa_2}{2}y_2^2. \tag{70}$$

The first-order condition under $y_1 + y_2 = 1/2$ is $\kappa_1 y_1 = \kappa_2 y_2$. Its solution is

$$y_1^\star = \frac{\kappa_2}{2(\kappa_1 + \kappa_2)}, \qquad y_2^\star = \frac{\kappa_1}{2(\kappa_1 + \kappa_2)}. \tag{71}$$

Both coordinates lie in $(0, 1/2)$. The solution is therefore feasible, and strict convexity makes it the unique optimum. Substituting the two curvature assignments and using $x_i = 1/2 + y_i$ gives

$$x^{\star,0} = (3/4 + d, 3/4 - d), \qquad x^{\star,1} = (3/4 - d, 3/4 + d), \tag{72}$$

as stated in Eq. (9). Each allocation can be selected by the instance-aware planner at the terminal decision. Let $\ell^\star_{\theta_i} = \ell_{\theta_i}(x^{\star,i})$. Evaluating the loss at the two instance-aware optima and at $x^{\mathrm{ref}} = (3/4, 3/4)$ gives

$$V^{(1)}_{\mathrm{full}} = \ell^\star_{\theta_0} = \ell^\star_{\theta_1} = \frac{3}{16} + \frac{\kappa_L\kappa_H}{8\Sigma}, \qquad V_{\mathrm{ref}} = \frac{3}{16} + \frac{\Sigma}{32}. \tag{73}$$

To optimize over transcript policies, parameterize the full feasible set by $x = (3/4 + z, 3/4 - z)$, $z \in [-1/4, 1/4]$. Completing the square in Eq. (70) gives

$$\ell_{\theta_0}(x) - \ell^\star_{\theta_0} = \frac{\Sigma}{2}(z - d)^2, \qquad \ell_{\theta_1}(x) - \ell^\star_{\theta_1} = \frac{\Sigma}{2}(z + d)^2. \tag{74}$$

Fix any admissible transcript policy π, and let Z be the displacement of its terminal allocation. By Lemma C.3, Z has the same law under both instances. Writing $\mathbb{E}$ for expectation under this common law and using the instance loss $L_\theta(\pi)$ from Eq. (28), we obtain

$$\frac{1}{2}\sum_{i=0}^{1}\big(L_{\theta_i}(\pi) - \ell^\star_{\theta_i}\big) = \frac{\Sigma}{4}\mathbb{E}\big[(Z - d)^2 + (Z + d)^2\big] = \frac{\Sigma}{2}\big(\mathbb{E}[Z^2] + d^2\big) \ge \frac{\Sigma d^2}{2} = G. \tag{75}$$

The reference attains equality by choosing $Z = 0$. Hence $V^{(1)}_{\mathrm{tr}} = V_{\mathrm{ref}}$, and $\mathrm{Gap}^{(1)}_{\mathrm{rec}} = V_{\mathrm{ref}} - V^{(1)}_{\mathrm{full}} = \frac{\Sigma d^2}{2} = \frac{(\kappa_H - \kappa_L)^2}{32\Sigma}$, which is Eq. (10). For the minimax comparison, the larger of the two expected excess losses is at least their average, so Eq. (75) gives a lower bound of G for every transcript policy. The reference incurs exactly G under each instance by Eq. (74). It therefore also attains the minimax bound. Together with Lemma C.3, this proves the theorem. □

For example, when $(\kappa_L, \kappa_H) = (1, 3)$, we have $\Sigma = 4$ and $d = 1/8$. The optimal allocations and values are $x^{\star,0} = (7/8, 5/8)$, $x^{\star,1} = (5/8, 7/8)$, $V_{\mathrm{full}}^{(1)} = \frac{9}{32}$, $V_{\mathrm{tr}}^{(1)} = V_{\mathrm{ref}} = \frac{5}{16}$, $G = \frac{1}{32}$.

C.3 Approximate Transcript Laws

Corollary C.4 (Approximate aliasing for the continuous pair)**.** *Consider the objective pair and feasible set in Theorem 4.1. Let W denote the record observed before the terminal allocation. Suppose the response interface induces laws satisfying $\mathrm{TV}(P_{\theta_0}^{\pi,W}, P_{\theta_1}^{\pi,W}) \le \varepsilon$ for every admissible transcript policy π. Then, under the uniform prior,*

$$\mathrm{Gap}_{\mathrm{rec}}^{(1)} \ge (1 - \varepsilon)G. \tag{76}$$

We now prove Corollary C.4. Keep the objectives and terminal feasible set fixed, and consider a response interface that may induce different transcript laws under the two instances. Fix an admissible transcript policy $\pi \in \Pi_{\mathrm{tr,feas}}^{(1)}(B)$, and let W be its complete record before the terminal allocation. Its record laws are $P_{\theta_i}^{\pi,W}$ for $i \in \{0, 1\}$. Under the uniform prior, the law of W is

$$\overline{P}^{\pi} = \frac{1}{2}\big(P_{\theta_0}^{\pi,W} + P_{\theta_1}^{\pi,W}\big). \tag{77}$$

Write $\eta(w) = \Pr(\theta = \theta_0 \mid W = w)$ for the posterior probability under this prior and fixed policy. At a record w, an allocation $x = (3/4 + z, 3/4 - z)$ has conditional expected excess loss

$$\frac{\Sigma}{2}\left[\eta(w)(z - d)^2 + (1 - \eta(w))(z + d)^2\right] = \frac{\Sigma}{2}\left[z - d(2\eta(w) - 1)\right]^2 + 2\Sigma d^2 \eta(w)(1 - \eta(w)). \tag{78}$$

The minimizing displacement is $z = d(2\eta(w) - 1)$. Since $\eta(w) \in [0, 1]$ and $d < 1/4$, it lies in $[-d, d]$ and is feasible. The minimum conditional excess loss is therefore $2\Sigma d^2 \eta(w)(1 - \eta(w))$. The bound holds for every feasible displacement, so it also holds after averaging over the policy's randomized action. Integrating over the record gives

$$\frac{1}{2}\sum_{i=0}^{1}\big(L_{\theta_i}(\pi) - \ell_{\theta_i}^{\star}\big) \ge 2\Sigma d^2 \int \eta(w)(1 - \eta(w))\, d\overline{P}^{\pi}(w). \tag{79}$$

We bound this integral using the binary testing error considered in Appendix B. For $\eta \in [0, 1]$,

$$\eta(1 - \eta) \ge \frac{1}{2}\min\{\eta, 1 - \eta\}. \tag{80}$$

Indeed, $\eta(1 - \eta)$ is the product of the smaller and larger of η and $1 - \eta$, and the larger is at least $1/2$. The quantity $\int \min\{\eta(w), 1 - \eta(w)\}\, d\overline{P}^{\pi}(w)$ is the optimal testing error under the uniform prior: at each record, choosing the more probable instance gives error $\min\{\eta(w), 1 - \eta(w)\}$. This error has a direct expression in total variation. Let r_i be the density of $P_{\theta_i}^{\pi,W}$ with respect to $\overline{P}^{\pi}$. Then $r_0 + r_1 = 2$ and $\eta = r_0/2$ almost surely under $\overline{P}^{\pi}$. It follows that

$$\begin{aligned}\int \min\{\eta, 1 - \eta\}\, d\overline{P}^{\pi} = \frac{1}{2}\int \min\{r_0, r_1\}\, d\overline{P}^{\pi} &= \frac{1}{2}\left(1 - \frac{1}{2}\int |r_0 - r_1|\, d\overline{P}^{\pi}\right) \\ &= \frac{1}{2}\left(1 - \mathrm{TV}\big(P_{\theta_0}^{\pi,W}, P_{\theta_1}^{\pi,W}\big)\right).\end{aligned} \tag{81}$$

Combining Eqs (79), (80), and (81) yields

$$\frac{1}{2}\sum_{i=0}^{1}\big(L_{\theta_i}(\pi) - \ell_{\theta_i}^{\star}\big) \ge \frac{\Sigma d^2}{2}\left(1 - \mathrm{TV}\big(P_{\theta_0}^{\pi,W}, P_{\theta_1}^{\pi,W}\big)\right) = G\left(1 - \mathrm{TV}\big(P_{\theta_0}^{\pi,W}, P_{\theta_1}^{\pi,W}\big)\right). \tag{82}$$

The assumption of Corollary C.4 bounds this total variation by ε for every admissible π. The instance-aware optimum is unchanged because the objectives and terminal action set are fixed. Taking the infimum over transcript policies therefore gives $\mathrm{Gap}_{\mathrm{rec}}^{(1)} \ge (1 - \varepsilon)G$ proving Eq. (76).

Table 5: Transcript aliasing under extensions of the response interface.

Query or returned information	Effect on the response	Aliasing
Ordinary price queries, with objective value and first derivative at the returned point	All added quantities are independent of κ.	Preserved
Exact optimizer responses for prices in $[-1/4, 1/4]$	The returned point is $1/4-\lambda$ for both curvatures.	Preserved
A price query at $\lambda_\star = -1/4 - \kappa_L/4$	The two optimizers in Eq. (83) differ.	Broken
A derivative query at $x \in (1/2, 1]$	The response $1/4 + \kappa(x - 1/2)$ depends on κ.	Broken
A proximal query with center $3/4$ and parameter $0 < \alpha < 1$	The optimizer in Eq. (84) depends on κ.	Broken

C.4 Extensions of the Response Interface

For this pair, a query can distinguish the instances only if its response depends on the private curvature assignment. We examine several extensions of the exact response rule below. For each comparison, the listed queries are admissible under the public resource and query limits. Table 5 summarizes their effect on transcript aliasing.

Additional information at returned points. For an ordinary price query $\lambda \in [-1/4, 1/4]$, the returned point is $x = 1/4 - \lambda \in [0, 1/2]$. At this point, $\varphi_\kappa(x) = \frac{\lambda^2}{2}$, $\varphi'_\kappa(x) = -\lambda$. The local objective value and first derivative are independent of κ. Adding either quantity to the response therefore preserves equality of the messages under the two instances. For $\lambda \in (-1/4, 1/4)$, the returned point lies in $(0, 1/2)$ and $\varphi''_\kappa(x) = 1$, so an interior second-derivative response is also aliased. At $\lambda = -1/4$, however, the response is $x = 1/2$, where the two one-sided curvatures are 1 and κ. The second derivative need not exist. We therefore make no closed-domain second-derivative claim. Lemma C.3 applies to adaptive policies using the well-defined additional observations above.

Any prescribed bounded price domain can be blind. The normalized construction extends to an arbitrary compact price interval $\Lambda = [\underline{\lambda}, \overline{\lambda}]$ with $\underline{\lambda} < \overline{\lambda}$. Let $c = (\underline{\lambda} + \overline{\lambda})/2$, $L = (\overline{\lambda} - \underline{\lambda})/2$, and $s = 4L$. On $[0, s]$, define $\widetilde{\varphi}_{\kappa,\Lambda}(x) = s^2\varphi_\kappa(x/s) - cx$. For every $\lambda \in \Lambda$, minimizing $\widetilde{\varphi}_{\kappa,\Lambda}(x) + \lambda x$ is equivalent, after setting $t = x/s$, to minimizing $\varphi_\kappa(t) + ((\lambda - c)/s)t$. Since $(\lambda - c)/s \in [-1/4, 1/4]$, Lemma C.2 gives the unique response $\widetilde{\mathcal{R}}_{\kappa,\Lambda}(\lambda) = \frac{s}{4} - (\lambda - c) \in [0, s/2]$, which is independent of κ. Scale the terminal feasible set and reference to $x_1 + x_2 = 3s/2$ and $(3s/4, 3s/4)$. The two optima become $(3s/4 \pm sd, 3s/4 \mp sd)$, and the exact Bayes and minimax gap is s^2G. Linear shifts do not change curvature, while this scaling preserves the smoothness and strong-convexity constants. Thus every predeclared bounded price domain admits a smooth strongly convex pair that it cannot distinguish, despite arbitrarily many adaptive queries.

Queries outside the blind range. For the normalized pair, prices that move the optimizer above $1/2$ can reveal the curvature. For example, consider the publicly determined price $\lambda_\star = -\frac{1}{4} - \frac{\kappa_L}{4}$. For $\kappa \in \{\kappa_L, \kappa_H\}$, the stationary point on the upper branch is $1/2 + \kappa_L/(4\kappa)$. It lies in $(1/2, 1)$, so strong convexity makes it the unique minimizer. Hence

$$\arg\min_{x\in[0,1]} \{\varphi_{\kappa_L}(x) + \lambda_\star x\} = \frac{3}{4}, \qquad \arg\min_{x\in[0,1]} \{\varphi_{\kappa_H}(x) + \lambda_\star x\} = \frac{1}{2} + \frac{\kappa_L}{4\kappa_H}. \tag{83}$$

Since $\kappa_L < \kappa_H$, the two responses differ. One exact query at $\lambda_\star$ to either party therefore identifies its curvature and distinguishes the two assignments.

Derivative queries on the upper branch. An exact derivative query at a publicly chosen $x \in (1/2, 1]$ returns $\varphi'_\kappa(x) = \frac{1}{4} + \kappa(x - 1/2)$. Since $x - 1/2 > 0$, the two possible curvatures give different responses. This includes the upper endpoint $x = 1$ and the reference coordinate $x = 3/4$. In particular, $\varphi'_\kappa(3/4) = \frac{1}{4} + \frac{\kappa}{4}$. Such a response identifies the queried party's curvature. At points in $[0, 1/2]$, the first derivative is instead independent of κ, as shown in Eq. (67).

Proximal queries. Consider a query with public center $v \in [0,1]$ and parameter $\alpha > 0$ that returns $\arg\min_{x\in[0,1]} \left\{\varphi_\kappa(x) + \frac{(x-v)^2}{2\alpha}\right\}$. Take $v = 3/4$ and $0 < \alpha < 1$. The first-order condition on the upper branch gives the optimizer

$$x_\kappa = \frac{1}{2} + \frac{1-\alpha}{4(1+\alpha\kappa)}. \tag{84}$$

This point lies in $(1/2, 3/4)$, and strong convexity makes it the unique minimizer. For fixed $\alpha \in (0,1)$, x_κ decreases strictly with κ, so one exact query distinguishes the two assignments.

D SEQUENTIAL K-SPARSE EXTENSION

This appendix proves Theorem D.1. We extend the construction in Appendix C by allowing the allocation to change through a sequence of bounded increments. The loss change produced by an increment depends on the current displacement, which is determined by earlier increments. The permitted price queries return the same responses under both private instances throughout the sequence. We retain the objectives in Eq. (4) and the two curvature assignments in Eq. (5). Recall that $0 < \kappa_L < \kappa_H$, $\Sigma = \kappa_L + \kappa_H$, $d = \frac{\kappa_H - \kappa_L}{4\Sigma} \in (0, 1/4)$, $G = \frac{\Sigma d^2}{2}$. We use the uniform prior on $\{\theta_0, \theta_1\}$ and suppress the fixed prior and resource budget in the Bayes value notation.

D.1 SEQUENTIAL ALLOCATIONS

Parameterize the feasible set in Eq. (6) by $x(z) = (3/4 + z, 3/4 - z)$, $z \in [-1/4, 1/4]$. Fix a horizon H, an intervention quota $K \in \{0, \dots, H\}$, and an increment bound $0 < \delta \le d$. The public coordination state is $S_t = z_t$, initialized at $z_0 = 0$, so the initial allocation is $x(0) = x^{\mathrm{ref}}$. The transcript also records the remaining resource budget B_t, query allowance Q_t, and intervention quota K_t, initially $(B_0, Q_0, K_0) = (B, Q, K)$. At each stage $t = 0, \dots, H-1$, the policy first selects a price query $q_t = (m, \lambda)$, with $m \in \{1, 2\}$ and $\lambda \in [-1/4, 1/4]$, or a null query $q^\emptyset$. A price query returns $\mathcal{R}_{m,\theta}(\lambda)$ from Eq. (7). A null query returns a fixed empty response and has zero cost. The budget B and query limit Q apply only to price queries. Query costs are accumulated as in Eq. (25). A price query is admissible when $Q_t > 0$ and its public cost does not exceed B_t. After the response, the policy chooses the coordination action $A_t = u_t$, where

$$u_t \in \mathcal{U}(z_t) = \{u \in [-\delta, \delta] : z_t + u \in [-1/4, 1/4]\}, \tag{85}$$

and $u_t = 0$ is required when $K_t = 0$. The public updates are

$$z_{t+1} = z_t + u_t, \qquad K_{t+1} = K_t - \mathbf{1}\{u_t \neq 0\}, \qquad Q_{t+1} = Q_t - \mathbf{1}\{q_t \neq q^\emptyset\}. \tag{86}$$

The remaining budget B_t is reduced by the public query cost. After a null query, the policy may still choose an admissible increment. The reference uses null queries and chooses $u_t = 0$ at every post-response transcript. The intervention indicator in Eq. (26) is therefore $Z_t^\pi = \mathbf{1}\{u_t \neq 0\}$, so at most K increments may be nonzero. Both policy classes satisfy these same constraints. The instance-aware planner may additionally use the private curvature assignment when choosing queries and increments. Every reached displacement lies in $[-1/4, 1/4]$, and the terminal allocation $x_H = x(z_H)$ belongs to $\mathcal{X}$. In this sequential model, we write $\ell_\theta(z)$ for the loss of allocation $x(z)$, as in Eq. (45):

$$\ell_\theta(z) = f_{1,\theta}(3/4 + z) + f_{2,\theta}(3/4 - z). \tag{87}$$

The notation $L_\theta(\pi)$ continues to denote the expected terminal loss of a policy, as defined in Eq. (28).

Theorem D.1 (Recoverability gap under bounded intervention reach)**.** *For every finite query limit Q and every budget B for which the protocol above is feasible, each admissible adaptive randomized transcript policy induces the same law for every public transcript prefix under θ_0 and θ_1. Let $r_K = \min\{K\delta, d\}$. Under the uniform prior, the recoverability gap is*

$$\mathrm{Gap}_{\mathrm{rec}}^{(K)} = V_{\mathrm{tr}}^{(K)} - V_{\mathrm{full}}^{(K)} = G_K = \frac{\Sigma}{2}\left[d^2 - (d - r_K)^2\right]. \tag{88}$$

Among admissible K-sparse transcript policies, the minimax excess loss relative to the K-sparse instance-aware optima also equals G_K. Thus $G_0 = 0$, the gap increases strictly with K while $K\delta < d$, and it equals G once $K\delta \ge d$.

D.2 Transcript Equivalence after Increments

Lemma D.2 (Transcript equivalence)**.** *For every admissible transcript policy, the laws of all pre-query and post-response transcript prefixes are identical under θ_0 and θ_1. Its terminal displacement therefore has the same law under both instances.*

Proof. Fix an admissible transcript policy π. The initial displacement and remaining resources are the same under both instances. Suppose the pre-query transcript laws agree at stage t. Applying the common query kernel gives the same joint law of the transcript and selected query. By Lemma C.2, every permitted price query returns $1/4 - \lambda$ under either curvature. This response rule does not depend on z_t, so the identity continues to hold after any admissible sequence of increments. Null queries also have the same response under both instances. The post-response transcript laws therefore agree. Applying the common action kernel gives the same joint law of the post-response transcript and u_t. The public displacement and resource updates depend only on this transcript and the selected increment, so the next pre-query transcript laws also agree. Induction over the H stages proves the prefix equivalence defined in Appendix A.7. Since z_H is part of the terminal transcript, its laws agree under the two instances. □

D.3 Reachable Allocations and Exact Values

Lemma D.3 (Reachable terminal displacements)**.** *The set of terminal displacements reachable from $z_0 = 0$ using at most K nonzero increments is*

$$\{z \in [-1/4, 1/4] : |z| \le K\delta\} = \big[-\min\{K\delta, 1/4\}, \min\{K\delta, 1/4\}\big]. \tag{89}$$

Proof. For any admissible sequence,

$$|z_H| = \left|\sum_{t=0}^{H-1} u_t\right| \le \sum_{t=0}^{H-1} |u_t| \le K\delta. \tag{90}$$

The state constraint also requires $|z_H| \le 1/4$. Conversely, fix z satisfying both bounds. The point $z = 0$ is reached by the reference. If $z \neq 0$, let $n = \lceil |z|/\delta \rceil \le K$. Take n increments in the direction of z: the first $n-1$ have magnitude δ, and the last has magnitude $|z| - (n-1)\delta \in (0, \delta]$. All intermediate displacements lie between 0 and z, so they satisfy the state constraint. Since $n \le K \le H$, these increments fit within the horizon. All remaining increments are zero. Null queries at every stage make this sequence feasible under the resource and query limits. □

Proof of Theorem D.1. Let $\vartheta = +1$ under θ_0 and $\vartheta = -1$ under θ_1. The instance-aware optimum over $\mathcal{X}$ has displacement ϑd. Its loss is the same under both instances. Write $\ell^\star = \ell_{\theta_0}(d) = \ell_{\theta_1}(-d) = \frac{3}{16} + \frac{\kappa_L \kappa_H}{8\Sigma}$ using Eq. (73). In the scalar state notation of Eq. (87), the identities in Eq. (74) become

$$\ell_{\theta_0}(z) - \ell^\star = \frac{\Sigma}{2}(z-d)^2, \qquad \ell_{\theta_1}(z) - \ell^\star = \frac{\Sigma}{2}(z+d)^2. \tag{91}$$

Instance-aware value. By Lemma D.3, the planner can choose any terminal displacement with $|z| \le K\delta$ and $|z| \le 1/4$. The loss is minimized by the reachable point closest to ϑd. Since $d < 1/4$, this point is

$$z_H^{\star,0} = r_K, \qquad z_H^{\star,1} = -r_K, \qquad r_K = \min\{K\delta, d\}. \tag{92}$$

Strict convexity makes these the unique optimal terminal displacements under the respective instances. The corresponding allocations are $x(r_K)$ and $x(-r_K)$. Every reachable terminal displacement has loss at least $\ell^\star + \Sigma(d - r_K)^2/2$, and the displacements in Eq. (92) attain this value. Consequently, this is also the optimum over randomized planners. The per-instance planner values defined in Eq. (32) satisfy

$$L_{\theta_0}^{\mathrm{full},\star} = L_{\theta_1}^{\mathrm{full},\star} = V_{\mathrm{full}}^{(K)} = \ell^\star + \frac{\Sigma}{2}(d - r_K)^2. \tag{93}$$

The quantity r_K is the optimal displacement magnitude. Once $K\delta \ge d$, we have $r_K = d$, and the planner reaches the optimum over $\mathcal{X}$. Its value is then $\ell^\star$.

Transcript value. Fix any admissible K-sparse transcript policy π, and let $Z = z_H$ be its terminal displacement. By Lemma D.2, Z has the same law under both instances. Write $\mathbb{E}$ for expectation under this common law. Since $L_{\theta_i}(\pi) = \mathbb{E}^{\pi}_{\theta_i}[\ell_{\theta_i}(Z)]$, the uniform-prior expected loss is

$$\frac{1}{2}\sum_{i=0}^{1} L_{\theta_i}(\pi) = \ell^\star + \frac{\Sigma}{4}\mathbb{E}\big[(Z-d)^2 + (Z+d)^2\big] = \ell^\star + \frac{\Sigma}{2}\big(\mathbb{E}[Z^2] + d^2\big) \geq \ell^\star + \frac{\Sigma d^2}{2}. \tag{94}$$

The reference attains equality with $Z = 0$. It follows that, for every intervention quota K,

$$V_{\mathrm{tr}}^{(K)} = V_{\mathrm{ref}} = \ell^\star + \frac{\Sigma d^2}{2} = \frac{3}{16} + \frac{\Sigma}{32}. \tag{95}$$

Subtracting Eq. (93) from Eq. (95) gives $\mathrm{Gap}_{\mathrm{rec}}^{(K)} = \frac{\Sigma d^2}{2} - \frac{\Sigma}{2}(d - r_K)^2 = G_K$, which is Eq. (88).

Minimax excess loss. Use the same K-sparse per-instance planner values in the minimax comparison of Eq. (33). For every admissible transcript policy,

$$\max_{i\in\{0,1\}} \big[L_{\theta_i}(\pi) - L^{\mathrm{full},\star}_{\theta_i}\big] \geq \frac{1}{2}\sum_{i=0}^{1}\big[L_{\theta_i}(\pi) - L^{\mathrm{full},\star}_{\theta_i}\big] = G_K + \frac{\Sigma}{2}\mathbb{E}[Z^2] \geq G_K. \tag{96}$$

The reference has loss $\ell^\star + \Sigma d^2/2$ on each instance, so its excess over the corresponding K-sparse planner value is exactly G_K on both. Thus $\mathcal{R}^{(K)}_{\mathrm{mm}}(\{\theta_0, \theta_1\}, B) = G_K$.

Dependence on the intervention quota. For fixed δ, the gap can be written as

$$G_K = \begin{cases} \Sigma d K\delta - \dfrac{\Sigma}{2}(K\delta)^2, & K\delta < d, \\ \dfrac{\Sigma d^2}{2} = G, & K\delta \geq d. \end{cases} \tag{97}$$

At $K = 0$, both policy classes must keep $z_H = 0$, so $G_0 = 0$. For $K < H$ with $K\delta < d$, we have $r_K < r_{K+1} \leq d$. Increasing the quota therefore reduces the planner excess from $\Sigma(d - r_K)^2/2$ to $\Sigma(d - r_{K+1})^2/2$. The transcript value in Eq. (95) remains unchanged, so $G_{K+1} > G_K$. Once $K\delta \geq d$, the planner reaches the optimum over $\mathcal{X}$, and the gap is G. In particular, if $\delta = d$ and $H \geq 1$, one intervention suffices and $G_1 = G$. Together with Lemma D.2, these calculations prove the theorem. □

D.4 Effect of Earlier Increments

Let $\vartheta = +1$ under θ_0 and $\vartheta = -1$ under θ_1. For a displacement z and an admissible increment u, subtracting the loss at $z + u$ from the loss at z in Eq. (91) gives

$$\ell_\theta(z) - \ell_\theta(z+u) = \frac{\Sigma}{2}\big[(z - \vartheta d)^2 - (z + u - \vartheta d)^2\big] = \Sigma\left(\vartheta d u - z u - \frac{u^2}{2}\right). \tag{98}$$

Its sign determines whether moving from z to $z + u$ decreases or increases the loss. The term $-\Sigma z u$ makes the loss difference depend on the current displacement. At stage t, $z_t = \sum_{s=0}^{t-1} u_s$, so earlier increments affect this difference through z_t. For example, if $z > 0$ and $u > 0$, the loss difference is smaller by $\Sigma z u$ than for the same increment from $z = 0$. Earlier nonzero increments also reduce the remaining quota K_t, which limits how many further nonzero increments the policy can take before the terminal allocation.

E Aliasing under a Representation

This appendix applies the testing bound to decisions based on a representation of the transcript and budget. It defines the restricted policy class and explains why retaining the public budget distinguishes cases with different fixed budgets.

Decision rules based on a representation. Consider the two-action setting of Theorem B.1, with $\mathcal{A} = \{a_0, a_1\}$ and two equiprobable cases indexed by $i \in \{0, 1\}$. Both actions are admissible in each case. Fix a common transcript policy for the interaction before this decision. Write $W = (\tau, B)$ for the decision record, where τ is the public transcript available at the decision and B is the public resource budget. Let P_i be the law of W in case i. Let ϕ be a measurable map from the decision record space to a representation space. The same map is used in both cases. Let g range over measurable probability kernels from the representation space to $\mathcal{A}$. The associated class of randomized decision rules is $\mathcal{G}_\phi = \{\pi_g : \pi_g(a \mid w) = g(a \mid \phi(w)), \quad a \in \mathcal{A}\}$. Each rule uses the same kernel g in both cases. Thus two records with the same representation lead to the same distribution over actions. Let P_i^ϕ denote the law of $\phi(W)$ in case i. For every measurable set D in the representation space, $P_i^\phi(D) = P_i\big(\{w : \phi(w) \in D\}\big)$, $i \in \{0, 1\}$.

Proposition E.1 (Aliasing under a representation)**.** *Let $J_i(a)$ be the loss of action a in case i, as in Theorem B.1. Suppose that, for some $\gamma > 0$,*

$$J_0(a_1) - J_0(a_0) \geq \gamma, \qquad J_1(a_0) - J_1(a_1) \geq \gamma. \tag{99}$$

For a decision rule $\pi \in \mathcal{G}_\phi$, its Bayes regret relative to choosing a_i in case i is

$$\mathrm{Reg}(\pi) = \frac{1}{2} \sum_{i=0}^{1} \int \sum_{a \in \mathcal{A}} \pi(a \mid w) \big[J_i(a) - J_i(a_i)\big] \, dP_i(w). \tag{100}$$

Every such rule satisfies

$$\mathrm{Reg}(\pi) \geq \frac{\gamma}{2} \big(1 - \mathrm{TV}(P_0^\phi, P_1^\phi)\big). \tag{101}$$

Proof. For $\pi = \pi_g$, the kernel g chooses the action from $\phi(W)$, whose law in case i is P_i^ϕ. Applying Eq. (65) with $\Delta = \gamma$ gives Eq. (101). □

Information retained by the representation. Under the uniform prior, the minimum error probability for distinguishing the two cases from $\phi(W)$ is $\big(1 - \mathrm{TV}(P_0^\phi, P_1^\phi)\big)/2$. The two cases have different preferred actions, and the margin conditions imply that choosing the wrong action incurs regret at least γ. The regret is therefore at least γ times this minimum error probability. In particular, if $P_0^\phi = P_1^\phi$, every rule in $\mathcal{G}_\phi$ has regret at least $\gamma/2$. The data-processing inequality in Eq. (64) gives $\mathrm{TV}(P_0^\phi, P_1^\phi) \leq \mathrm{TV}(P_0, P_1)$. Thus identical laws of the full decision record remain identical after applying ϕ. A representation can also discard information that distinguishes the two cases. For example, a constant map gives $P_0^\phi = P_1^\phi$ even when $\mathrm{TV}(P_0, P_1) = 1$.

Retaining the public budget. The budget is an input to $\phi(\tau, B)$. We say that the representation retains the budget if B can be recovered from $\phi(W)$ by the same measurable function in both cases. This holds, for example, when the budget is a coordinate of the representation. Suppose the two cases have distinct fixed budget values b_0 and b_1, so that $B = b_i$ almost surely in case i and $b_0 \neq b_1$. If the representation retains B, let D_0 be the measurable set of representation values for which the recovered budget equals b_0. Then $P_0^\phi(D_0) = 1$, $P_1^\phi(D_0) = 0$, $\mathrm{TV}(P_0^\phi, P_1^\phi) = 1$. The retained budget therefore distinguishes these two cases, and the lower bound in Eq. (101) is zero. Different budgets can favor different actions. Retaining B allows a rule in $\mathcal{G}_\phi$ to assign different action probabilities at different budgets. The cross-budget formulation in Appendix A.8 also allows query and action rules to depend on the observed budget.

F Decision Certificates

This appendix proves the results in Section 5. We specify the certificate model at a terminal decision, derive the necessary separation of the augmented record laws, and prove the sufficiency of a locally computed one-bit certificate. We also calculate the remaining gap under independent bit flips and give a counting result for finite instance families.

F.1 The Augmented Policy Class

We use the augmented policy class $\Pi^{(1)}_{\mathrm{tr}+C,\mathrm{feas}}(B, B_C)$ in Eq. (42). The intervention quota is $K = 1$, and the certificate is received after the ordinary responses and before the terminal coordination action. Let W be the ordinary record available at this decision. Since no earlier certificates have been sent, $W = \tau_t^+$ at the certificate stage, as in Eq. (38). A designated party m uses its local private information θ_m and W to generate a message through a measurable kernel, $C_m \sim K_m(\cdot \mid \theta_m, W)$. This is the single-stage form of Eq. (39). Each sender uses fresh local randomness, independent of the coordinator's and environment's random sources. If several parties send, their messages are conditionally independent given θ and W, and C lists them in the public sender order. The coordinator chooses its terminal action by a randomized rule based on (W, C). The public record determines which parties send. The encoded message C uses at most B_C bits, as in Eq. (41). The ordinary interaction follows the response rule and constraints of the original protocol. A fixed certificate scheme Γ specifies the public transmission rules and local message kernels. Following Eq. (43), write $V^{(1)}_{\mathrm{tr}+C}(\nu; B, B_C \mid \Gamma)$ for the optimal Bayes loss when these message rules are fixed and the coordinator's query and action rules are optimized. Its remaining recoverability gap is $V^{(1)}_{\mathrm{tr}+C}(\nu; B, B_C \mid \Gamma) - V^{(1)}_{\mathrm{full}}(\nu, B)$. Optimizing also over the allowed certificate schemes gives $V^{(1)}_{\mathrm{tr}+C}(\nu; B, B_C)$. Eq. (44) gives

$$V^{(1)}_{\mathrm{full}}(\nu, B) \le V^{(1)}_{\mathrm{tr}+C}(\nu; B, B_C) \le V^{(1)}_{\mathrm{tr}}(\nu, B). \tag{102}$$

For the continuous pair of Theorem 4.1, we retain the public constants $0 < \kappa_L < \kappa_H$, $\Sigma = \kappa_L + \kappa_H$, $d = \frac{\kappa_H - \kappa_L}{4\Sigma} \in (0, 1/4)$, $G = \frac{\Sigma d^2}{2}$. The prior ν is uniform on $\{\theta_0, \theta_1\}$.

F.2 Necessary Separation

Corollary F.1 (Certificate necessity)**.** *Let π be an admissible augmented transcript policy, and let $P_i^{\pi,W,C}$ denote the law of its observed record (W, C) under θ_i. In the two-action setting of Theorem B.1, with the margins in Eq.* (46)*, if* $\mathrm{Reg}(\pi) \le \epsilon$ *for* $0 \le \epsilon < \Delta/2$*, then*

$$\mathrm{TV}\big(P_0^{\pi,W,C}, P_1^{\pi,W,C}\big) \ge 1 - \frac{2\epsilon}{\Delta}. \tag{103}$$

For the continuous pair of Theorem 4.1, under the uniform prior ν, if

$$\mathbb{E}_{\theta\sim\nu} L_\theta(\pi) - V^{(1)}_{\mathrm{full}}(\nu, B) \le \epsilon, \qquad 0 \le \epsilon < G, \tag{104}$$

then

$$\mathrm{TV}\big(P_0^{\pi,W,C}, P_1^{\pi,W,C}\big) \ge 1 - \frac{\epsilon}{G}. \tag{105}$$

Proof. In the two-action problem, the coordinator's action rule uses the augmented record (W, C). Applying Theorem B.1 to its two laws gives

$$\mathrm{Reg}(\pi) \ge \frac{\Delta}{2}\left(1 - \mathrm{TV}\big(P_0^{\pi,W,C}, P_1^{\pi,W,C}\big)\right). \tag{106}$$

Combining this inequality with $\mathrm{Reg}(\pi) \le \epsilon$ proves Eq. (103). For the continuous pair, fix π and define the mixture law

$$\overline{P}^{\pi,W,C} = \frac{1}{2}\big(P_0^{\pi,W,C} + P_1^{\pi,W,C}\big). \tag{107}$$

Let $\eta(w, c) = \Pr(\theta = \theta_0 \mid W = w, C = c)$ be the posterior under this policy and the uniform prior. Parameterize the terminal allocation by $x = (3/4 + z, 3/4 - z)$, with $z \in [-1/4, 1/4]$. The two excess losses in Eq. (74) give the following conditional expected excess at (w, c):

$$\begin{aligned} &\frac{\Sigma}{2}\left[\eta(w,c)(z-d)^2 + (1-\eta(w,c))(z+d)^2\right] \\ &\qquad = \frac{\Sigma}{2}\left[z - d(2\eta(w,c) - 1)\right]^2 + 2\Sigma d^2 \eta(w,c)(1-\eta(w,c)). \end{aligned} \tag{108}$$

The minimizing displacement is $d(2\eta(w,c)-1)$. It lies in $[-d,d]$ and is feasible because $d<1/4$. The minimum conditional excess is therefore $2\Sigma d^2\eta(w,c)(1-\eta(w,c))$. A randomized allocation averages the losses of its possible choices and satisfies the same lower bound. Integrating yields

$$\mathbb{E}_{\theta\sim\nu}L_\theta(\pi)-V^{(1)}_{\mathrm{full}}(\nu,B)\geq 2\Sigma d^2\int \eta(w,c)(1-\eta(w,c))\,d\overline{P}^{\pi,W,C}(w,c). \tag{109}$$

By Eq. (80), $\eta(1-\eta)\geq\frac{1}{2}\min\{\eta,1-\eta\}$. The testing-error identity in Eq. (81), applied to (W,C), gives

$$\int \min\{\eta(w,c),1-\eta(w,c)\}\,d\overline{P}^{\pi,W,C}(w,c)=\frac{1}{2}\left(1-\mathrm{TV}\big(P_0^{\pi,W,C},P_1^{\pi,W,C}\big)\right). \tag{110}$$

Consequently,

$$\mathbb{E}_{\theta\sim\nu}L_\theta(\pi)-V^{(1)}_{\mathrm{full}}(\nu,B)\geq G\left(1-\mathrm{TV}\big(P_0^{\pi,W,C},P_1^{\pi,W,C}\big)\right). \tag{111}$$

Combining this with Eq. (104) proves Eq. (105). This gives the continuous-pair condition. □

Corollary F.2 (One-bit necessity)**.** *For the continuous pair and uniform prior of Theorem 4.1, suppose an augmented transcript policy π attains the instance-aware Bayes value using a deterministic certificate $C=c(\theta,W)$. Here c is the combined map induced by the local certificate rules. Let P^W be the common law of W under the two instances. Then*

$$c(\theta_0,w)\neq c(\theta_1,w)\quad \textit{for } P^W\textit{-almost every } w. \tag{112}$$

The message alphabet contains at least two symbols, so a fixed-length encoding requires at least one bit.

Proof. The certificate is sent after the ordinary interaction. Lemma C.3 therefore gives the common law P^W for the record before it is received. Let $E=\{w: c(\theta_0,w)=c(\theta_1,w)\}$. For any measurable set D of augmented records, the contributions from E cancel, and hence

$$P_0^{\pi,W,C}(D)-P_1^{\pi,W,C}(D)=\int_{E^c}\left[\mathbf{1}_D(w,c(\theta_0,w))-\mathbf{1}_D(w,c(\theta_1,w))\right]dP^W(w). \tag{113}$$

The integrand has absolute value at most one. Taking the supremum over D gives

$$\mathrm{TV}\big(P_0^{\pi,W,C},P_1^{\pi,W,C}\big)\leq 1-P^W(E). \tag{114}$$

Attaining the instance-aware Bayes value requires total variation equal to one by Corollary F.1 with $\epsilon=0$. Thus $P^W(E)=0$, proving Eq. (112). A fixed-length binary encoding of these distinct messages requires at least one bit. The lower bound is attained by the local bit. □

Randomized certificates. Let π be an admissible policy using a randomized certificate with finite message alphabet $\mathcal{C}$ in the same hard pair. Write $k_i(c\mid w)=\mathbb{P}^\pi_{\theta_i}(C=c\mid W=w)$. Since the marginal law of W is P^W in both instances,

$$1-\mathrm{TV}\big(P_0^{\pi,W,C},P_1^{\pi,W,C}\big)=\int\sum_{c\in\mathcal{C}}\min\{k_0(c\mid w),k_1(c\mid w)\}\,dP^W(w). \tag{115}$$

Indeed, since the two joint laws have the same marginal P^W,

$$\mathrm{TV}\big(P_0^{\pi,W,C},P_1^{\pi,W,C}\big)=\frac{1}{2}\int\sum_{c\in\mathcal{C}}|k_0(c\mid w)-k_1(c\mid w)|\,dP^W(w). \tag{116}$$

Using $\min\{a,b\}=(a+b-|a-b|)/2$ and $\sum_{c\in\mathcal{C}}k_i(c\mid w)=1$ gives Eq. (115). If π attains the instance-aware Bayes value, Corollary F.1 requires the total variation to equal one. The integral in Eq. (115) is therefore zero. Since every summand is nonnegative and $\mathcal{C}$ is finite, for P^W-almost every w, no message has positive conditional probability under both instances.

F.3 A Locally Computed One-Bit Certificate

The curvature values κ_L and κ_H are public. The two instances differ in how these values are assigned to the parties. By Eq. (67), $\varphi'_\kappa(3/4) = \frac{1}{4} + \frac{\kappa}{4}$. The two possible marginal costs at the reference are therefore the public values in Eq. (16), $m_L = \frac{1}{4} + \frac{\kappa_L}{4}$, $m_H = \frac{1}{4} + \frac{\kappa_H}{4}$. Following Eq. (17), party 1 sends

$$C_1 = \mathbf{1}\{f_1'(3/4) = m_L\}. \tag{117}$$

This message is computed from its local objective and the public value m_L. It is sent once before the terminal allocation, with $B_C = 1$.

Theorem F.3 (One-bit certificate sufficiency, restated)**.** *For the continuous pair of Theorem 4.1, the allocation*

$$x_1 = \frac{3}{4} + (2C_1 - 1)d, \qquad x_2 = \frac{3}{2} - x_1 \tag{118}$$

is the unique instance-aware optimum under either instance. Under the uniform prior,

$$V^{(1)}_{\mathrm{tr}+C}(\nu; B, 1) = V^{(1)}_{\mathrm{full}}(\nu, B). \tag{119}$$

Proof. Under θ_0, party 1 has curvature κ_L, so $C_1 = 1$ and Eq. (118) gives $(3/4+d, 3/4-d) = x^{\star,0}$. Under θ_1, party 1 has curvature κ_H, so $C_1 = 0$ and the allocation is $(3/4-d, 3/4+d) = x^{\star,1}$. These are the unique optima in Eq. (9). Both allocations belong to the terminal feasible set $\mathcal{X}$ in Eq. (6) and can be selected with one intervention. The policy can use null queries before the certificate, satisfying the ordinary resource and query limits. The certificate has two possible values and uses one bit. Thus an admissible augmented policy attains the instance-aware minimum on each instance. The value ordering in Eq. (102) then gives Eq. (119), proving Theorem 5.1. □

The bit also determines the direction in which to change the allocation from the reference. Along the feasible direction $(1, -1)$, the directional derivative is

$$f_1'(3/4) - f_2'(3/4) = \begin{cases} m_L - m_H < 0, & \theta = \theta_0, \\ m_H - m_L > 0, & \theta = \theta_1. \end{cases} \tag{120}$$

When $C_1 = 1$, this derivative is negative and the optimal displacement is $+d$. When $C_1 = 0$, it is positive and the optimal displacement is $-d$. The magnitude d is public. The bit specifies which of these two displacements to use.

F.4 A Noisy Certificate

Let Γ_p be the scheme that sends the bit in Eq. (117) and flips it with known probability $p \in [0, 1/2]$, independently of the private instance and the ordinary interaction. Write $\widetilde{C}_1$ for the received bit, so the observed certificate is $C = \widetilde{C}_1$.

Corollary F.4 (Noisy certificate)**.** *Under the uniform prior, the Bayes-optimal displacement for Γ_p is*

$$z = (2\widetilde{C}_1 - 1)(1 - 2p)d. \tag{121}$$

The corresponding remaining recoverability gap is

$$V^{(1)}_{\mathrm{tr}+C}(\nu; B, 1 \mid \Gamma_p) - V^{(1)}_{\mathrm{full}}(\nu, B) = 4p(1-p)G. \tag{122}$$

Proof. For any admissible ordinary query policy, the record W has the same law under the two instances by Lemma C.3. The posterior before the certificate is therefore uniform. Since $C_1 = 1$ under θ_0 and $C_1 = 0$ under θ_1, Bayes' rule gives

$$\Pr(\theta = \theta_0 \mid W, \widetilde{C}_1 = 1) = 1 - p, \qquad \Pr(\theta = \theta_0 \mid W, \widetilde{C}_1 = 0) = p, \tag{123}$$

almost surely under the uniform-prior mixture. The posterior is therefore determined by the received bit and the known flip probability p. For a posterior probability η of θ_0, Eq. (108) is minimized at $z = d(2\eta - 1)$. Substituting the two posterior probabilities in Eq. (123) gives Eq. (121). This displacement lies in $[-d, d]$ and is feasible. For either received bit, the minimum conditional excess loss is $2\Sigma d^2 \eta(1-\eta) = 2\Sigma d^2 p(1-p) = 4p(1-p)G$. The nonnegative square in Eq. (108) shows

that no randomized allocation has smaller conditional expected loss. The lower bound holds for every policy with the scheme Γ_p. Using null queries and the allocation in Eq. (121) attains the bound with at most one intervention and one certificate bit. Taking the infimum over policies with the fixed scheme Γ_p proves Eq. (122). □

At $p = 0$, the received bit gives the instance-aware optimum. At $p = 1/2$, the optimal displacement is zero, giving the reference allocation and remaining gap G.

F.5 Finite Instance Families

Consider a single decision with a finite private-instance family Θ, a common feasible action set $\mathcal{A}$, and a measurable loss $J_\theta(a)$ under instance θ. Fix an admissible ordinary transcript policy before the decision, and suppose the resulting record W has the same law P^W for every $\theta \in \Theta$. Assume that each J_θ has a unique minimizer $a^\star_\theta$. Group the instances according to their optimal action: two instances belong to the same class if their minimizers coincide. Let N be the number of classes, and let $a^\star_1, \dots, a^\star_N$ be their distinct optimal actions. Write $j(\theta) \in \{1, \dots, N\}$ for the class containing θ, so that $a^\star_\theta = a^\star_{j(\theta)}$.

Corollary F.5 (Certificates for a finite instance family)**.** *Suppose a deterministic certificate $C = c(\theta, W)$, followed by a possibly randomized decision rule based on (W, C), attains the minimum expected loss on every instance. Then its message alphabet $\mathcal{C}$ satisfies*

$$|\mathcal{C}| \geq N. \tag{124}$$

A fixed-length binary encoding therefore requires at least $\lceil \log_2 N \rceil$ bits. If the actions $a^\star_1, \dots, a^\star_N$ are publicly specified and a designated party m can compute $j(\theta)$ as a measurable function of (θ_m, W) almost surely under every instance, then a certificate of $\lceil \log_2 N \rceil$ bits is sufficient.

Proof. On each instance, the excess loss $J_\theta(A) - J_\theta(a^\star_\theta)$ is nonnegative. Its expectation is zero, so it is zero almost surely. Uniqueness of the minimizer therefore gives $A = a^\star_\theta$ almost surely. Thus, for P^W-almost every w, the decision rule assigns probability one to $a^\star_\theta$ after observing $(w, c(\theta, w))$. Choose one representative $\theta^{(j)}$ from each class. Since there are finitely many representatives and their record laws are all P^W, there is a common set of records of P^W-probability one on which this property holds for every representative. Fix w in this set. If representatives from two distinct classes sent the same message, the coordinator would use the same action distribution at that augmented record. It could not assign probability one to both of their different optimal actions. Consequently, the N messages $c(\theta^{(1)}, w), \dots, c(\theta^{(N)}, w)$ are all distinct. This proves Eq. (124). A fixed-length code of b bits has at most 2^b codewords, so $2^b \geq N$ and $b \geq \lceil \log_2 N \rceil$. For sufficiency, the designated party computes $j(\theta)$ from its local information and W and sends its fixed-length binary encoding. The coordinator decodes the index and selects the public action $a^\star_{j(\theta)}$. This action is feasible and optimal on the corresponding instance. The construction uses the stated number of bits and attains the minimum loss on every instance. □

For the continuous hard pair, there are two classes. Party 1 can determine its class from its local curvature, and the two optimal allocations are specified by the public constants. The certificate in Eq. (117) therefore satisfies the local computation condition and attains the one-bit bound.

F.6 List-Free Approximate Optimality

Assume a valid subgradient $g \in \partial F(x)$ and a known lower curvature bound $\mu > 0$. The function may be nonsmooth. What is required is the strong-subgradient inequality

$$F(y) \geq F(x) + g^\top (y - x) + \frac{\mu}{2}\|y - x\|^2, \qquad x, y \in \mathcal{X}. \tag{125}$$

For a separable sum of local convex objectives, stacking one valid local subgradient at each block gives such a g. If the local approximation errors have Euclidean bounds δ_m, their stacked error is at most $\delta = (\sum_m \delta_m^2)^{1/2}$.

Proof of Proposition 5.2. Choose $n \in N_{\mathcal{X}}(x)$ closest to $-\widehat{g}$. Such a vector exists because the normal cone is closed and nonempty. For $y \in \mathcal{X}$, $n^\top(y-x) \le 0$, so the strong-subgradient inequality implies

$$F(y) \ge F(x) + (g+n)^\top(y-x) + \frac{\mu}{2}\|y-x\|^2 \ge F(x) - \frac{\|g+n\|^2}{2\mu}. \tag{126}$$

The triangle inequality gives $\|g+n\| \le \|\widehat{g}+n\| + \delta = \widehat{r}(x) + \delta$. Minimizing over $y \in \mathcal{X}$ proves Eq. (18). □

The same proof permits any computed $n \in N_{\mathcal{X}}(x)$, using $\|\widehat{g}+n\|$ instead of the minimum distance. Thus an approximate normal-cone projection is usable as long as its returned normal vector is feasible. The error bound must come from the response computation or encoding, not from comparing the received message with an evaluator-only gradient. This is an application of the standard strong-convexity inequality (Boyd & Vandenberghe, 2004), used here to specify an additional response interface.

Public aggregate constraints. For $\mathcal{X} = \{y : 0 \le y_{mh} \le U_{mh},\ \sum_m y_{mh} = b_h\}$, a normal has the form $n_{mh} = \lambda_h + s_{mh}$. Here $s_{mh} = 0$ at an interior coordinate, $s_{mh} \le 0$ at the lower bound, and $s_{mh} \ge 0$ at the upper bound. At a fixed λ_h, the residual components are $\widehat{g}_{mh} + \lambda_h$ in the interior, their negative parts at a lower bound, and their positive parts at an upper bound. Minimizing their squared norm is a one-dimensional convex problem for each hour. This calculation uses only public constraints and the received message.

Value-function curvature and errors. In the Appendix G task, let $V_m(y_m)$ be the minimum resource cost at aggregate allocation y_m. Weighted Cauchy–Schwarz gives $\sum_r c_{mrh}(x_{mrh} - x'_{mrh})^2 \ge \frac{(y_{mh}-y'_{mh})^2}{\sum_r c_{mrh}^{-1}}$. Applying strong convexity to the two local minimizing resource schedules shows that V_m is strongly convex with constant at least $\min_h(\sum_r c_{mrh}^{-1})^{-1}$. Since $c_{mrh} \ge 0.35 m_r$ and $m_r = 0.65 + 0.3(r+1)$ is public, a valid common public constant is $\mu_{\mathrm{pub}} = \min_m \frac{0.35}{\sum_{r=0}^{R_m-1} m_r^{-1}}$. The public bound need not equal the realized curvature. An equality multiplier determines a subgradient of V_m. The Appendix K bisection bracket contains a valid multiplier even when the solution is an interval. Its final midpoint is within the bracket half-width of a valid multiplier. Quantization to a grid of width Δ adds at most $\Delta/2$ per coordinate. Each party sends the Euclidean norm of these coordinatewise error bounds, along with its subgradient vector. The certificate evaluations in Appendix N use the same public μ_{pub} and count these additional messages and 50-step equality solves. That appendix also specifies the floating-point safeguards and results.

G Regional Coordination Task and Evaluation Protocol

This appendix describes the generated resource-allocation task used in the response-protocol, sparse-intervention, and record-compatibility experiments of Section 6. Appendix M separately defines the controlled perturbation task. The reference-anchoring experiment uses another configuration to study objective deterioration relative to the reference policy.

G.1 Data Source and Participating Countries

We use hourly records from the Open Power System Data (OPSD) Time Series package, version 2020-10-06 (Open Power System Data, 2020). These records include country-level actual load, day-ahead load forecasts, and wind or solar generation from the ENTSO-E Transparency Platform, all measured in MW. The dataset contains 50,401 hourly timestamps, from 2014-12-31 23:00 to 2020-09-30 23:00 UTC. We use these records to construct private allocation costs, resource capacities, and public demand. Each experiment has $M = 8$ parties, each representing one country. Party m has R_m virtual resources. Table 6 gives the country assigned to each party, the number of virtual resources, the renewable-generation signal, and the public query costs for both experimental configurations. Timestamps are converted to UTC and grouped by calendar date. We retain only days with exactly 24 distinct hourly observations and finite values for every required signal. We use these complete days without interpolation or imputation. For each configuration, the normalization parameters are

Table 6: Party configurations and public query costs. A query to party m costs c_m^{comm} communication units and Jc_m^{loc} local-work units when it uses J solver iterations.

m	Code	Country	Renewable signal	R_m	c_m^{comm}	c_m^{loc}
Main experiments						
1	PT	Portugal	Onshore wind	5	1	20
2	SI	Slovenia	Solar	5	3	20
3	FI	Finland	Onshore wind	8	4	8
4	CH	Switzerland	Solar	8	3	16
5	SK	Slovakia	Solar	3	1	3
6	NO	Norway	Onshore wind	5	2	10
7	SE	Sweden	Onshore wind	5	1	15
8	LV	Latvia	Onshore wind	4	4	4
Reference-anchoring experiment						
1	DK	Denmark	Wind	2	1	2
2	PL	Poland	Wind	3	1	3
3	IT	Italy	Wind	4	1	4
4	ES	Spain	Solar	5	1	5
5	GR	Greece	Wind	6	1	6
6	RO	Romania	Wind	3	1	3
7	DE	Germany	Solar	4	1	4
8	CZ	Czechia	Wind	5	1	5

estimated using only its training interval. Appendix G.6 specifies the chronological intervals and the role of each evaluation subset.

G.2 Private Allocation Costs and Public Demand

Each allocation problem spans one day. Let $h \in \{1, \dots, 24\}$ index its hours and $r \in \{1, \dots, R_m\}$ index the resources of party m. The hour index is separate from the interaction stage t and protocol horizon H in Section 3. The allocation variables are expressed in normalized units, and the input time series are measured in MW.

Normalized signals. Write L_{mh}, $\widehat{L}_{mh}$, and G_{mh} for actual load, day-ahead load forecast, and renewable generation. From the training interval, define the party's load scale $\ell_m = \max\{\mathrm{median}\,|L_m|, 1\}$, where the median is over that party's retained training observations. The forecast-error and renewable-generation ratios are $\epsilon_{mh} = \frac{|L_{mh} - \widehat{L}_{mh}|}{\ell_m}$, $\rho_{mh} = \frac{G_{mh}}{\ell_m}$. For a training vector v, set $s(v) = \max\left\{1.4826\,\mathrm{median}\,|v - \mathrm{median}(v)|, 10^{-6}\right\}$. The standardized signals are

$$e_{mh} = \mathrm{clip}_{[-6,6]}\left(\frac{\epsilon_{mh} - \mathrm{median}(\epsilon_m)}{s(\epsilon_m)}\right), \qquad g_{mh} = \mathrm{clip}_{[-6,6]}\left(\frac{\rho_{mh} - \mathrm{median}(\rho_m)}{s(\rho_m)}\right). \quad (127)$$

Here $\mathrm{clip}_{[a,b]}(v) = \min\{b, \max\{a, v\}\}$. Let ϵ_m and ρ_m collect these ratios over party m's retained training observations. We estimate all medians and scales on the training interval and use them unchanged on subsequent intervals.

Resource capacities and preferences. Let $\sigma(v) = (1 + \exp(-v))^{-1}$. For each party m and resource r, we draw $\varphi_{mr} \sim \mathrm{Unif}[-\pi, \pi]$, $\xi_{mr} \sim \mathcal{N}(0, 0.2^2)$. Write $m_0 = m - 1$, $r_0 = r - 1$, and $h_0 = h - 1$, and define

$$q_{mrh} = \xi_{mr} - 0.12 r_0 + 0.35 \sin\left(\frac{h_0 + \varphi_{mr} + m_0}{2 + 0.3 r_0}\right), \qquad a_{mrh} = \frac{\exp(q_{mrh})}{\sum_{r'=1}^{R_m} \exp(q_{mr'h})}. \quad (128)$$

The shares satisfy $a_{mrh} > 0$ and $\sum_r a_{mrh} = 1$. The party's hourly capacity and preferred allocation are $U_{mh} = 0.75 + 0.2\sigma(g_{mh})$, $P_{mh} = U_{mh}\sigma(1.1 e_{mh} + 0.25 g_{mh})$. These quantities are distributed among the resources using the same shares $\bar{x}_{mrh} = U_{mh} a_{mrh}$, $p_{mrh} = P_{mh} a_{mrh}$. Thus the capacities sum to U_{mh} and the preferred allocations sum to P_{mh}. In particular, $0 < p_{mrh} < \bar{x}_{mrh}$.

Private costs. For $\mu_{mr} = 0.65 + 0.3r$, define the quadratic coefficient $c_{mrh} = [0.35 + 1.4\sigma(e_{mh})]\mu_{mr}$. The absolute-deviation coefficient is determined by

$$b_{mrh} = 0.8e_{mh} - 0.45g_{mh} + 0.55\sin\left(\frac{h_0 + \varphi_{mr} + 3m_0}{2.5 + 0.2r_0}\right) + 0.35r_0,$$
$$\omega_{mrh} = [0.04 + 0.28\sigma(b_{mrh})]\mu_{mr}. \tag{129}$$

For a resource allocation $x_m = (x_{mrh})_{r,h}$ satisfying $0 \le x_{mrh} \le \bar{x}_{mrh}$, party m has objective

$$f_m(x_m) = \sum_{r=1}^{R_m}\sum_{h=1}^{24}\left[\frac{c_{mrh}}{2}(x_{mrh} - p_{mrh})^2 + \omega_{mrh}|x_{mrh} - p_{mrh}|\right]. \tag{130}$$

Both coefficients are positive. The quadratic term makes f_m strongly convex, and the absolute-deviation term introduces a kink at $x_{mrh} = p_{mrh}$. Without the shared balance constraint, the unique minimizer is $x_{mrh} = p_{mrh}$ for every resource and hour. The realized cost parameters and resource capacities belong to the party's local private information θ_m. The coordinator observes their effect through the response interface described below. The instance-aware evaluator uses the full collection of private parameters to evaluate candidate continuations and compute the centralized optimum.

Public demand and the shared constraint. Let $d_h^{\mathrm{pref}} = \sum_{m=1}^{M} P_{mh}$, $\bar{U}_h = \sum_{m=1}^{M} U_{mh}$, $s_h = \frac{1}{M}\sum_{m=1}^{M}|e_{mh}|$. Here d_h^{pref} is the total preferred allocation, $\bar{U}_h$ is the total available capacity, and s_h is the mean absolute value of the standardized forecast-error signal e_{mh} across parties. We use s_h to define the stress demand $d_h^{\mathrm{stress}} = \bar{U}_h[0.2 + 0.68\sigma(1.35s_h)]$. At conflict level γ, public demand is $d_h(\gamma) = \mathrm{clip}_{[0.02\bar{U}_h,\, 0.98\bar{U}_h]}\left((1-\gamma)d_h^{\mathrm{pref}} + \gamma d_h^{\mathrm{stress}}\right)$. Before clipping, γ controls the weight on the stress demand relative to the total preferred allocation. For the chosen conflict level, write $d_h = d_h(\gamma)$. The allocation must satisfy hourly balance:

$$\sum_{m=1}^{M}\sum_{r=1}^{R_m} x_{mrh} = d_h, \qquad h = 1, \ldots, 24. \tag{131}$$

When $d_h \neq d_h^{\mathrm{pref}}$, meeting demand requires some resources to move away from their preferred allocations. The allocation problem determines how to distribute this adjustment across resources with different private costs. Because $\sum_{m,r}\bar{x}_{mrh} = \bar{U}_h$ and $d_h \in [0, \bar{U}_h]$, the box constraints and hourly balance have a common feasible allocation. The constraints are separable across hours. Separable local costs with a shared balance constraint follow the distributed energy-management formulation of Kraning et al. (2014). The reference-anchoring experiment uses five sets of resource parameters and conflict levels $\gamma \in \{0.5, 0.75, 1.0\}$. For a fixed seed, φ_{mr} and ξ_{mr} remain unchanged across days. The hourly records determine the daily capacities, preferred allocations, and cost coefficients.

G.3 Response Interface and Resource Budgets

Public state and private local problem. The coordinator state contains a feasible aggregate allocation $z^{(t)} \in \mathbb{R}^{M\times 24}$, a scaled dual variable $u^{(t)}$ of the same shape, and one cached aggregate response $\widetilde{y}_m^{(t)}$ per party. The aggregate bounds $U_{mh} = \sum_r \bar{x}_{mrh}$, aggregate preferred totals $P_{mh} = \sum_r p_{mrh}$, public demand, query costs, and penalty ρ are protocol data. Resource-level capacities, preferred allocations, and cost coefficients remain local to the party. Initially, $u^{(0)} = 0$, $\widetilde{y}^{(0)} = P$, and $z^{(0)}$ is the Euclidean projection of P onto the public aggregate feasible set defined below.

A query selects a party and a local-work level, $q = (m, J)$, $m \in \{1, \ldots, M\}$, $J \in \{2, 5, 10, 25, 50\}$. At the current public state, party m approximately solves

$$\min_{0\le x_{mrh}\le \bar{x}_{mrh}} f_m(x_m) + \frac{\rho}{2}\sum_{h=1}^{24}\left(\sum_{r=1}^{R_m} x_{mrh} - z_{mh}^{(t)} + u_{mh}^{(t)}\right)^2. \tag{132}$$

The 24 hours separate. For one hour, let $v_{mh}^{(t)} = z_{mh}^{(t)} - u_{mh}^{(t)}$. Given a scalar multiplier ξ, the resource allocation is

$$x_{mrh}(\xi) = \mathrm{clip}_{[0,\bar{x}_{mrh}]}\left(p_{mrh} + \frac{\mathrm{sign}(-\xi)[|\xi| - \omega_{mrh}]_+}{c_{mrh}}\right). \tag{133}$$

The implementation applies exactly J midpoint-bisection updates to the monotone equation

$$r_{mh}(\xi) = \xi - \rho\left(\sum_r x_{mrh}(\xi) - v_{mh}^{(t)}\right) = 0. \tag{134}$$

Write $U_{mh} = \sum_r \bar{x}_{mrh}$ and define

$$b_{mh} = \max_r \left\{\left|c_{mrh}(p_{mrh} - \bar{x}_{mrh}) - \omega_{mrh}\right|, \left|c_{mrh}p_{mrh} + \omega_{mrh}\right|\right\} + \rho \max\{U_{mh}, |v_{mh}^{(t)}|, 1\}. \tag{135}$$

Each query starts from $[\xi_{mh}^-, \xi_{mh}^+] = [-b_{mh}, b_{mh}]$. At an update, it evaluates the midpoint $\xi = (\xi_{mh}^- + \xi_{mh}^+)/2$. If $r_{mh}(\xi) > 0$, it sets $\xi_{mh}^+ = \xi$, and otherwise sets $\xi_{mh}^- = \xi$. After J updates, the solver evaluates Eq. (133) at the final bracket midpoint. The bracket is recomputed for every query, and no previous local iterate is used as a warm start. All hours use the same declared J. If $\widehat{x}_{mrh}^{(t)}$ is the resulting allocation, the response is only the 24-vector

$$y_{mh}^{(t)} = \sum_{r=1}^{R_m} \widehat{x}_{mrh}^{(t)}, \qquad h = 1, \ldots, 24. \tag{136}$$

Thus the ordinary response Y_t of Section 3.1 has the same dimension for every allowed J. Increasing J changes response accuracy, not message dimension.

Cache, projection, and dual update. Only the queried party's cached vector is replaced: $\widetilde{y}_m^{(t+1/2)} = y_m^{(t)}$, while $\widetilde{y}_j^{(t+1/2)} = \widetilde{y}_j^{(t)}$ for $j \neq m$. The coordinator then applies

$$\begin{aligned} z^{(t+1)} &= \arg\min_{z \in \mathcal{Z}} \frac{1}{2}\left\| z - \left(\widetilde{y}^{(t+1/2)} + u^{(t)}\right)\right\|_F^2, \\ u^{(t+1)} &= u^{(t)} + \widetilde{y}^{(t+1/2)} - z^{(t+1)}, \end{aligned} \tag{137}$$

where

$$\mathcal{Z} = \left\{ z : 0 \le z_{mh} \le U_{mh},\ \sum_{m=1}^{M} z_{mh} = d_h \text{ for every } h \right\}. \tag{138}$$

The projection separates by hour and is computed with 60 bisection steps on its scalar balance multiplier. These updates, including the cached stale responses of unqueried parties, are deterministic functions of the preceding public state and the received response. A synchronous batch, used for initialization, evaluates all parties from the same pre-batch (z, u) and applies Eq. (137) once after replacing the batch of cached responses. The public state therefore consists of the cache, z, u, remaining resources, and query history. Private parameters enter a transcript policy only through the returned aggregate response. The next-query choice follows the chronology in Appendix A.2: after response Y_t, the coordination action A_t selects the party and local-work level for q_{t+1}. An intervention changes this choice relative to the reference action evaluated at the current transcript. It is counted when the choice is made, and the query is charged when it is issued.

Initialization and query costs. Each trajectory begins with one synchronous query to every party, all at $J = 2$ and all evaluated from the common initial state described above. These responses are included in the transcript, and the queries count toward total resource use, as in Appendix A.2. For a query to party m, the two costs are

$$c^{\mathrm{qry}}(m, J) = \left(c_m^{\mathrm{comm}},\ J c_m^{\mathrm{loc}}\right). \tag{139}$$

The constants in Table 6 give an initialization cost of $(19, 192)$ in the main experiments and $(8, 64)$ in the reference-anchoring experiment, in communication and local-work units. Following Appendix A.5, write the total resource budget as $B = (B^{\mathrm{comm}}, B^{\mathrm{loc}})$ and the budget remaining before query q_t as $B_t = (B_t^{\mathrm{comm}}, B_t^{\mathrm{loc}})$. A query to party m with local-work level J is affordable when

$$c_m^{\mathrm{comm}} \le B_t^{\mathrm{comm}}, \qquad J c_m^{\mathrm{loc}} \le B_t^{\mathrm{loc}}. \tag{140}$$

When selecting q_{t+1} after response Y_t, we apply these inequalities using the budget remaining after q_t has been charged.

Budget settings. In the main experiments, we vary the communication budget over $B^{\mathrm{comm}} \in \{27, 35, 51, 75, 107\}$ while fixing $B^{\mathrm{loc}} = 100{,}000$, and vary the local-work budget over $B^{\mathrm{loc}} \in \{256, 384, 576, 864, 1296\}$ while fixing $B^{\mathrm{comm}} = 160$. Candidate queries must satisfy both resource limits in Eq. (140). The six budgets with post-initialization decision positions under the reference are communication budgets $35, 51, 75, 107$ and local-work budgets $864, 1296$. Under the reference schedule, the other four budget settings have no decision positions after initialization. They are included in the aggregate results over all ten budget settings, with zero intervention opportunity under that reference. The reference-anchoring experiment uses communication budgets $\{8, 12, 16, 24, 32\}$ and local-work budgets $\{64, 128, 256, 512, 1024\}$. Its configuration and evaluation are described separately in Appendix I.

G.4 Reference Policies and Instance-Aware Evaluators

Reference policy. The reference is selected on the reference-validation interval from a finite class of static schedules. Each schedule fixes a party order, a local-work level, and a coordination penalty. The selected schedule is fixed before evaluation and is used subject to the public resource limits. After an intervention, the reference continuation uses the same selected schedule from the resulting state and remaining budget. Subsequent actions are compared with the reference choice at the current transcript, as defined in Appendix A.5. The round-robin order repeats parties $1, \ldots, 8$. The capacity and inverse-cost orders are deterministic 64-contact weighted cycles, using respectively the public capacity shares and weights proportional to $1/c_m^{\mathrm{loc}}$. For normalized weights w_m, integer contact counts are obtained from $\lfloor 64 w_m \rfloor$ with at least one contact per party and adjusted to sum to 64. At cycle position t, the next party maximizes $(t+1)w_m - n_m(t)$ among parties with contacts remaining, where $n_m(t)$ is its number of preceding contacts. The synchronous pattern repeatedly queries all eight parties from a common pre-batch state. Table 7 gives the schedule selected for every budget.

Table 7: Selected static reference schedules. C and L denote communication and local-work sweeps. J is the bisection count per contacted party.

Budget	Party-order pattern	J	ρ
C-27	synchronous	10	0.25
C-35	round robin	10	0.25
C-51	capacity	25	0.25
C-75	round robin	10	0.25
C-107	round robin	10	0.50
L-256	capacity	10	0.25
L-384	capacity	25	0.25
L-576	capacity	25	0.25
L-864	inverse cost	5	0.25
L-1296	inverse cost	5	0.25

Table 8: Reference and instance-aware searches. Each continuation uses the same state transitions and query costs.

Comparison	Query choices and continuation
Static reference	Follow the selected schedule from the current state and budget.
Best single intervention	Choose one reference position and one affordable replacement; continue with the reference from the resulting state.
Sequential, quota K	Choose the best improving position and replacement, observe its response, and repeat the search from the reached state, at most K times.
Sequential, no intervention quota	Repeat the same improving-replacement procedure until no improving feasible candidate remains.
Open-loop segment	Replace one reference segment by two to four queries fixed in advance, matching both resource costs over that segment.
Exact joint search, quota K	Search complete sequences with at most K replacements in the specified candidate class and compare their terminal values.

Instance-aware intervention searches. The evaluators use private objectives to compare the terminal regret of candidate continuations. Table 8 distinguishes the searches used in the opportunity experiments. Their search rules and complete results appear in Appendix H. Among candidates satisfying the raw balance violation condition, the evaluator selects the replacement that yields the largest reduction in terminal regret. After following the corresponding reference prefix and executing this replacement, it observes the response and repeats the search from the resulting state. Without an intervention quota, the procedure continues until no improving feasible replacement remains, giving $R_{\infty}^{\mathrm{seq}}$ in Eq. (156). The joint search evaluates complete intervention sequences. It retains sequences whose terminal raw balance violation is at most 10^{-8} above that of the reference trajectory from the same starting state and budget. The exact joint search minimizes terminal regret over complete sequences of candidate queries with at most K interventions, subject to the resource limits and the raw balance violation condition. The open-loop comparison replaces one reference segment with two to four queries fixed in advance. The replacement has the same total communication and local-work costs as the reference segment. After the segment, the reference schedule continues from the resulting state.

Evaluator-selected alternatives. Public query costs and the remaining budget determine the affordable alternatives at each decision state. For each affordable alternative, the evaluator computes terminal regret under the reference continuation from the resulting state. Candidates are retained when their terminal raw balance violation is no more than 10^{-8} above that of the reference continuation from the current state. Among these candidates, the evaluator selects the alternative with the smallest terminal regret. The transcript rule uses only the public transcript to decide whether to keep the reference query or use the evaluator-selected alternative, as in Section 6.3. The prediction studies in Section 6.3 use the evaluator to label comparisons between keeping and replacing the reference query on decision records. Appendix J specifies the continuation for each target and separates record-level prediction from recursively executed rules. The privileged inputs in the representation comparison are used for offline prediction.

G.5 Terminal Allocations and Evaluation Measures

Aggregate terminal action and private value functions. The public terminal action is the final projected aggregate allocation $z^{\pi} \in \mathcal{Z}$. No resource-level allocation is returned to the coordinator and no additional terminal message is exchanged. To score this public action, the evaluator uses each party's private value function

$$F_{m,\theta_m}(z_m) = \min_{x_m} \left\{ f_m(x_m) : 0 \le x_{mrh} \le \bar{x}_{mrh},\ \sum_r x_{mrh} = z_{mh}\ \forall h \right\}. \tag{141}$$

The equality-constrained minimizer is computed offline by an 80-step bisection on one multiplier per hour. This disaggregation is evaluator-side scoring, not a public state transition or an action available to the transcript policy. It is permitted because the terminal loss ℓ_θ in Section 3.1 may depend on the private instance, whereas the public transition kernel does not.

The centralized optimal value for an instance is

$$\begin{aligned} f^{\star} = \min_{(x_m)_{m=1}^{M}} \quad & \sum_{m=1}^{M} f_m(x_m) \\ \text{s.t.} \quad & 0 \le x_{mrh} \le \bar{x}_{mrh} \quad \text{for all } m, r, h, \\ & \sum_{m,r} x_{mrh} = d_h \quad \text{for all } h. \end{aligned} \tag{142}$$

This value is computed offline using the complete instance. It serves as the reference point for normalized terminal regret. Equivalently,

$$f^{\star} = \min_{z \in \mathcal{Z}} \sum_{m=1}^{M} F_{m,\theta_m}(z_m), \tag{143}$$

and the terminal objective of policy π is $F_\theta(z^\pi) = \sum_m F_{m,\theta_m}(z_m^\pi)$. The normalization scale is

$$s_{\mathrm{obj}} = \max\left\{ \sum_{m,r,h} \left(\frac{c_{mrh}}{2} p_{mrh}^2 + \omega_{mrh} |p_{mrh}| \right), 10^{-6} \right\}. \tag{144}$$

The sum is the objective at the zero allocation. For a policy π, define

$$R_\pi = \frac{F_\theta(z^\pi) - f^\star}{s_{\mathrm{obj}}}. \tag{145}$$

The private instance fixes $f^\star$ and s_{obj} for all policy comparisons on that instance. The terminal loss in the formal model can therefore be taken as this normalized regret. The value $V_{\mathrm{full}}^{(K)}(\nu, B)$ in Eq. (1) is the optimal expected normalized regret over instance-aware policies satisfying the protocol, resource budget, and intervention quota.

Raw balance violation. Let $\widetilde{y}_{mh}$ be the most recent aggregate response of party m for hour h before the final projection is read out. Define

$$V_{\mathrm{raw}} = \frac{\sum_{h=1}^{24} \left| \sum_{m=1}^{M} \widetilde{y}_{mh} - d_h \right|}{\sum_{m,r,h} \bar{x}_{mrh}}. \tag{146}$$

This measures the hourly imbalance in the cached responses before projection. The projected terminal action z^π satisfies the public balance constraint up to the numerical projection tolerance. The reference-relative difference is

$$\Delta V_{\mathrm{raw}}(\pi) = V_{\mathrm{raw}}(\pi) - V_{\mathrm{raw}}(\pi_{\mathrm{ref}}). \tag{147}$$

A negative difference means that the responses have less imbalance than under the reference. In the intervention searches, a candidate satisfies the comparison condition when its terminal raw balance violation is at most 10^{-8} above that of the reference continuation used by the search.

Regret reduction and intervention frequency. Write R_{ref} for reference regret. When $R_{\mathrm{ref}} > 0$, the relative regret reduction is

$$\mathrm{RRI}(\pi) = \frac{R_{\mathrm{ref}} - R_\pi}{R_{\mathrm{ref}}}. \tag{148}$$

An intervention is a departure from the reference action at the current transcript, as in Eq. (26). For the sequential opportunity study, the reported position rate divides the number of such changes by the number of reference positions. When a trigger rule is executed along a trajectory, the trigger rate divides the number of executed triggers by the number of eligible positions reached by the rule. These are the denominators used in Figure 3 and Table 19, respectively.

Sequential and policy capture. The sequential capture in Eq. (156) compares R_K^{seq} with the improvement obtained by continuing the same sequential evaluator until no improving feasible replacement remains. Its denominator is $R_{\mathrm{ref}} - R_\infty^{\mathrm{seq}}$. For a rule π executed with at most K interventions, policy capture uses the exact K-intervention evaluator for that study. Let R_K^{eval} denote its terminal regret. When $R_{\mathrm{ref}} > R_K^{\mathrm{eval}}$, define

$$\mathrm{PolicyCapture}_K(\pi) = \frac{R_{\mathrm{ref}} - R_\pi}{R_{\mathrm{ref}} - R_K^{\mathrm{eval}}}. \tag{149}$$

The evaluator's candidate actions and continuation rule are those of the corresponding experiment in Appendix J. In Section 6.3, the evaluator supplies the alternative query and the transcript rule decides whether to use it. To measure improvement beyond the best single intervention, let $R_K^{\mathrm{glob},\star}$ denote the regret returned by the joint search in Appendix H. When $R_1^{\mathrm{glob},\star} > R_3^{\mathrm{glob},\star}$, define

$$\mathrm{MultiCapture}(\pi) = \frac{R_1^{\mathrm{glob},\star} - R_\pi}{R_1^{\mathrm{glob},\star} - R_3^{\mathrm{glob},\star}}. \tag{150}$$

This quantity is negative when the policy has higher regret than the best single intervention. Capture measures are reported without clipping.

Aggregation across runs. Aggregate regret reductions and capture values are ratios of sums over the same evaluation runs. For example, for a set $\mathcal{U}$ of runs,

$$\mathrm{Capture}_K^{\mathrm{seq}}(\mathcal{U}) = \frac{\sum_{u\in\mathcal{U}}(R_{\mathrm{ref},u} - R_{K,u}^{\mathrm{seq}})}{\sum_{u\in\mathcal{U}}(R_{\mathrm{ref},u} - R_{\infty,u}^{\mathrm{seq}})}. \tag{151}$$

The ratio is defined when its aggregate denominator is positive. Runs with zero evaluator improvement contribute zero to that denominator and remain in the evaluation set. The same aggregation convention applies to Eqs (148), (149), and (150).

Net capture on decision records. The transfer experiments in Section 6.3 evaluate scores on a fixed collection of labeled decision records. For record v and remaining quota k_v, define

$$A_v = R_{\mathrm{keep}}(v, k_v) - R_{\mathrm{trigger}}(v, k_v). \tag{152}$$

Both alternatives start from the same decision state and use the same continuation rule after the current action. Their exact continuations depend on the experiment and are specified in Appendix J. A positive A_v favors the replacement. For a record index set $\mathcal{D}$ and a selected subset $I \subseteq \mathcal{D}$, net capture is

$$\mathrm{NC}(I) = \frac{\sum_{v\in I} A_v}{\sum_{v\in\mathcal{D}}[A_v]_+}, \qquad [a]_+ = \max\{a, 0\}. \tag{153}$$

This definition requires positive total advantage in the denominator. Negative advantages among the selected records reduce the numerator. The evaluation records and denominator remain fixed when comparing scores within the same evaluation setting. For a score g, let $\mathrm{Top}_q(g)$ select the $\mathrm{round}(q|\mathcal{D}|)$ highest-scoring records, using the fixed tie-breaking order. Ranking net capture is

$$\mathrm{RNC}(g) = \frac{1}{5} \sum_{q\in\{0.02,0.05,0.10,0.20,0.30\}} \mathrm{NC}\big(\mathrm{Top}_q(g)\big). \tag{154}$$

Figure 4 reports this average over five selection rates. The matched-rate comparison in Table 24 uses one common selection rate per budget. For the position-ranking result in Section 6.3, we select the three highest-scoring positions within each reference trajectory. The reported fraction is the sum of positive advantages at these positions divided by the sum of positive advantages over all positions in the evaluated reference trajectories.

Objective deterioration in the anchoring experiment. For a scenario w, let $F_{\mathrm{raw}}(\pi; w)$ be the normalized private-evaluator objective at the terminal cached aggregate responses, before the public aggregate projection. Define

$$U_w = \mathbf{1}\left\{F_{\mathrm{raw}}(\pi; w) > F_{\mathrm{raw}}(\pi_{\mathrm{ref}}; w)\right\}. \tag{155}$$

The empirical rate of objective deterioration is the mean of U_w. For each day, we average these indicators over five sets of resource parameters, three conflict levels, and ten budgets. The resulting daily loss is the average over 150 scenarios and lies in $[0, 1]$. Learn-then-Test (LTT) (Angelopoulos et al., 2025) uses these day-level losses with a risk target of 0.05. Appendix I describes the calibration and test evaluation. This risk measure concerns objective deterioration. V_{raw} measures imbalance, and the reported regret percentiles use the objective at the projected aggregate state.

G.6 Chronological Splits and Evaluation Units

Table 9 gives the full chronological partition for each task configuration. The intervals include their start date and exclude their end date. Day counts apply after the complete-day filter in Appendix G.1. The main opportunity and transcript-learning studies use fixed, disjoint subsets of the 352-day policy-development interval. The opportunity study uses 24 evaluation days, and the joint search uses eight separate evaluation days. The prediction studies use separate days for fitting, model selection, and evaluation, with the partitions for each study given in Appendix J. In transfer evaluations, the held-out budget is excluded from model fitting and from model and threshold selection. The fixed-record opportunity, joint-search, and predictor comparisons use dates in 2018. These predictor splits are held out by calendar day but are not posed as chronological forecasting experiments. The online residual-timing comparison instead uses 48 later dates from January through June 2019. Appendix J lists both sets explicitly.

Table 9: Chronological data partitions. All dates are in UTC.

Data role	Interval [start, end)	Days
Main experiments		
Training	2015-01-01 to 2017-01-01	669
Reference validation	2017-01-01 to 2017-07-01	168
Opportunity validation	2017-07-01 to 2018-01-01	171
Policy development	2018-01-01 to 2019-01-01	352
Risk calibration	2019-01-01 to 2019-07-01	174
Reserved test	2019-07-01 to 2020-01-01	180
Reference-anchoring experiment		
Training	2017-01-01 to 2019-01-01	722
Reference validation	2019-01-01 to 2019-04-01	90
Policy validation	2019-04-01 to 2019-07-01	88
Risk calibration	2019-07-01 to 2020-01-01	183
Anchoring test	2020-01-01 to 2020-10-01	272

Fixed evaluation dates. The 24 sparse-opportunity dates are `2018-01-04, 01-06, 01-08, 01-13, 01-22, 01-23, 03-13, 03-18, 04-02, 04-20, 04-26, 06-10, 06-20, 06-27, 08-20, 09-22, 10-14, 10-18, 11-03, 11-18, 11-19, 12-04, 12-11, 12-17`, where an omitted year is 2018. The eight disjoint joint-search dates are `2018-02-27, 05-06, 06-14, 08-17, 10-07, 11-24, 12-06, 12-10`.

Table 10: Evaluation sets and reported comparisons.

Analysis	Evaluation set	Comparison unit
Sparse opportunity	24 days; 2,400 runs	Aggregate regret and capture; consistency across objective seeds, conflict levels, and budgets.
Joint intervention search	Eight separate days; 480 runs	Complete terminal trajectories at each intervention quota.
Transcript rules and prediction	Separate days for fitting, model selection, and evaluation	Recursive policy outcomes or labeled-record predictions, according to the experiment; budgets define the transfer folds.
Reference anchoring	272 test days; 40,800 scenarios	Paired policy comparisons on the same scenario, and day-level averages for the risk analysis.

Table 10 lists the evaluation units for each analysis. Each run in the intervention study is defined by a day, a set of resource parameters, a conflict level, and a resource budget. For the main opportunity results, we examine whether regret reductions remain positive across parameter sets and conflict levels. The reference-anchoring test contains $40{,}800$ scenarios. Policy comparisons are paired by day, resource parameters, conflict level, and budget. Confidence intervals for mean regret differences use 50,000 hierarchical bootstrap replicates, resampling parameter sets first and then contiguous seven-day blocks. The test risk analysis uses one average loss per day, giving 272 daily losses. The complete tail summaries and calibration procedure are in Appendix I.

H Sparse Intervention Value

This appendix describes the intervention searches and additional results for Section 6.3. Appendix G describes the main coordination task used here and specifies the reference policy, query costs, and evaluation measures. Each run fixes a day, a set of resource parameters, a conflict level, and a resource budget. For the results below, let R_{ref} be reference regret, R_K^{seq} the regret after at most K sequential interventions, and R_∞^{seq} the regret when the same evaluator continues until no improving feasible replacement remains. For a positive denominator, sequential capture is

$$\mathrm{Capture}_K^{\mathrm{seq}} = \frac{R_{\mathrm{ref}} - R_K^{\mathrm{seq}}}{R_{\mathrm{ref}} - R_\infty^{\mathrm{seq}}}. \tag{156}$$

Aggregate capture is a ratio of sums over the same runs.

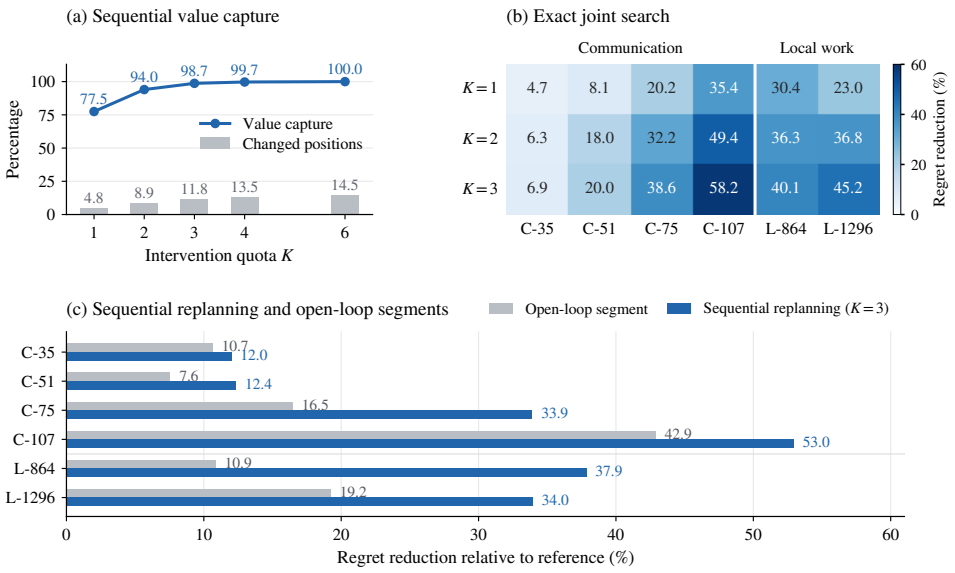


Figure 3: Sparse intervention value. (a) Sequential value capture and the fraction of reference positions changed. (b) Regret reductions from exact joint search on an independent evaluation set. (c) Regret reductions from sequential replanning and open-loop query segments. C and L denote communication and local-work budgets.

H.1 Sequential Intervention Search

Candidate queries and decision positions. A query specifies a party and a local-work level, $q = (m, J)$, with $m \in \{1, \dots, M\}$ and $J \in \{2, 5, 10, 25, 50\}$. We use $M = 8$ parties. At a decision following response Y_t, the reference action selects the next query q_{t+1}^{ref}. A replacement selects an affordable query $q \neq q_{t+1}^{\mathrm{ref}}$. Affordability is checked against both resource limits in Eq. (140), using the budget remaining after the current query has been charged. The intervention is counted when the replacement is selected, and the query cost is charged when it is issued, as in Appendix A.2.

Evaluating a single replacement. Let τ denote the post-response transcript from which a search starts. It determines the current coordination state, the reference schedule's position, and the remaining resource budget. Following the reference from this transcript gives terminal regret $R_{\mathrm{ref}}(\tau)$ and terminal raw balance violation $V_{\mathrm{raw}}^{\mathrm{ref}}(\tau)$. For each remaining decision position t on this reference continuation, including the current position, the evaluator follows the reference prefix to that position and considers the affordable replacement queries there. For a candidate q, it evaluates the response to that query, updates the coordination state, and follows the reference from the resulting transcript and remaining budget. Write $R(\tau; t, q)$ and $V_{\mathrm{raw}}(\tau; t, q)$ for the terminal regret and terminal raw balance violation of this complete continuation. The evaluator uses the private instance to compute the terminal objective. A candidate is retained when

$$V_{\mathrm{raw}}(\tau; t, q) \leq V_{\mathrm{raw}}^{\mathrm{ref}}(\tau) + 10^{-8}. \tag{157}$$

Among retained candidates, the evaluator selects the position and query with the smallest terminal regret. It executes the replacement only if this regret is smaller than $R_{\mathrm{ref}}(\tau)$. Raw balance violation is measured from the terminal cached responses before public projection. Regret is evaluated at the projected aggregate state, following Appendix G.5.

Continuing after an intervention. Once a replacement is selected, the trajectory follows the reference up to its decision position and issues the selected query. After receiving its response, the evaluator starts the next search from the resulting transcript and remaining budget. It recomputes the reference continuation and candidate values from this point. The raw balance violation of this reference continuation supplies the comparison value in Eq. (157). The evaluator makes at most K interventions. When the quota is exhausted or no retained candidate improves terminal regret, the

reference completes the trajectory. The resulting terminal regret is R_K^{seq}. Without an intervention quota, the same procedure continues until no improving feasible replacement remains. Its terminal regret is R_∞^{seq}. Eq. (156) measures the fraction of this procedure's total regret reduction obtained with at most K interventions. Its denominator is $R_{\mathrm{ref}} - R_\infty^{\mathrm{seq}}$, where R_{ref} is the regret of the reference trajectory from initialization.

H.2 Value from a Small Number of Interventions

The sequential evaluation uses 24 days, five sets of resource parameters, two conflict levels, and ten budget settings, giving 2,400 runs. The data partitions and parameter settings are described in Appendix G.6 and Appendix G.2, respectively. Aggregate regret reductions and capture values use the ratios of sums defined in Appendix G.5. Table 11 gives the results as the intervention quota

Table 11: Sequential intervention results over all 2,400 runs. Capture is defined in Eq. (156). Position rate is the fraction of reference positions changed. Regret reduction is relative to the reference, and ΔV_{raw} is defined in Eq. (147).

K	Capture (%)	Position rate (%)	Regret reduction (%)	Mean ΔV_{raw}
1	77.48	4.80	7.15	-0.00313
2	94.01	8.93	8.67	-0.00405
3	98.70	11.84	9.10	-0.00436
4	99.71	13.48	9.20	-0.00442
6	99.99	14.49	9.22	-0.00445

increases. With at most one, two, and three interventions, the evaluator captures 77.48%, 94.01%, and 98.70% of the improvement obtained by the same sequential procedure without an intervention quota. At $K = 3$, the evaluator changes 11.84% of reference positions and reduces regret by 9.10% relative to the reference. Increasing the quota from three to six raises capture to 99.99% and regret reduction to 9.22%.

Date-block uncertainty. The 2,400 runs share 24 calendar profiles and are therefore not treated as independent time samples. We resample the 24 dates with replacement, moving all parameter, conflict, and budget runs from a sampled date together, and recompute each ratio of sums. With 20,000 replicates and fixed seed 20270921, the percentile 95% intervals at $K = 3$ are $[98.27, 99.03]\%$ for capture and $[7.91, 10.44]\%$ for regret reduction. Table 12 reports the $K = 3$ results separately for

Table 12: Sequential results at $K = 3$ by resource parameters and conflict level. The parameter sets are identified by their random seeds. Each parameter row aggregates both conflict levels and all budgets. Each conflict row aggregates all parameter sets and budgets. The final column is the percentage of runs with at least one intervention.

Group	Value	Regret reduction (%)	Runs with an intervention (%)
Resource parameters	1709	9.61	59.79
	1721	8.65	60.00
	1723	8.73	60.00
	1733	9.46	60.00
	1741	9.10	60.00
Conflict level	0.75	9.36	59.92
	1.00	8.87	60.00

each set of resource parameters and each conflict level. The aggregate regret reduction is positive in every listed group. Six budget settings permit decisions after initialization under the reference schedule. Table 13 shows that, at $K = 3$, capture ranges from 97.14% to 99.68% across these settings. Regret reductions relative to the reference range from 12.02% to 52.95%. The other four budget settings have no such decision positions. The evaluator makes no interventions in those runs, which remain included in the aggregate results over all 2,400 runs, as described in Appendix G.3.

Table 13: Sequential results at $K = 3$ for the six budgets that permit decisions after initialization under the reference.

Budget	Capture (%)	Regret reduction (%)	Position rate (%)
Communication 35	99.68	12.02	27.44
Communication 51	99.52	12.36	16.73
Communication 75	97.14	33.92	12.57
Communication 107	97.19	52.95	7.93
Local work 864	99.26	37.86	13.57
Local work 1296	97.48	33.96	9.32

H.3 Open-Loop Query Segments

Replacing a reference segment. We compare sequential replanning with a single open-loop segment of two to four queries. A candidate specifies a consecutive window of the reference trajectory and a sequence of queries to execute in its place. The replacement queries are chosen before the segment begins. During execution, each party responds at the current coordination state, and the coordinator updates that state after each response. Once the segment is complete, the reference continues from the resulting transcript and remaining budget. Each replacement segment has the same total communication and local-work costs as the reference segment it replaces. Let $\mathcal{W}$ be the set of indices of the reference queries in that segment, and write the replacement queries as $q^{(j)} = (m_j, J_j)$ for $j = 1, \dots, n$, with $n \in \{2, 3, 4\}$. Using the vector query cost in Eq. (139), we require

$$\sum_{j=1}^{n} c^{\mathrm{qry}}(q^{(j)}) = \sum_{i \in \mathcal{W}} c^{\mathrm{qry}}(q_i^{\mathrm{ref}}), \tag{158}$$

where $c^{\mathrm{qry}}(q) = c^{\mathrm{qry}}(m, J)$ for $q = (m, J)$. The equality applies to both components of the cost vector. Terminal regret is evaluated after the segment and its reference continuation have completed. Candidate segments change the queried parties, the local-work levels, or the query order, individually or in combination. Each candidate must satisfy the resource limits and Eq. (158).

Table 14: Open-loop segment results with matched communication and local-work costs. Positive values in the third column favor the segment over the repeated single-query baseline. All entries are percentages.

Budget	Gain vs. reference	Gain vs. repeated	Trace mismatch	Segment coverage
Communication 35	10.65	6.48	43.99	49.58
Communication 51	7.57	4.36	26.38	28.40
Communication 75	16.48	−2.12	14.44	14.82
Communication 107	42.90	6.74	8.85	9.14
Local work 864	10.88	−2.51	22.14	24.11
Local work 1296	19.25	0.76	12.40	13.13

Results. Table 14 compares the open-loop segments with the reference and the repeated single-query baseline across the six budget settings. The aggregate trace mismatch is 16.19%. In 26.00% of runs, the segment reassigns at least 10% of the local work after initialization. The segment improves on the repeated single-query baseline in four of the six budget settings. The largest reported gain in this comparison is 6.74%. The $K = 3$ sequential planner achieves greater regret reduction than the open-loop segment in all six budgets, as shown in Figure 3c and Tables 13 and 14. The open-loop queries are fixed in advance, whereas sequential replanning chooses later replacements from the updated state. The two search classes also differ in where replacements may be placed, so this comparison measures the declared search classes rather than an isolated feedback effect.

H.4 Exact Joint Intervention Search

Jointly selecting interventions. We next evaluate the regret attainable when up to K interventions are selected jointly. The search starts from the transcript after the fixed initialization and considers

complete continuations of the coordination protocol. At each decision, it considers the reference query and the affordable alternatives $q = (m, J)$ defined in Appendix G.3. For each candidate continuation, the evaluator computes the responses, coordination states, and remaining resources along that continuation. Following the reference choice leaves the intervention quota unchanged. A different query consumes one intervention when selected, according to Eq. (26). Its communication and local-work costs are charged when the query is issued. The reference choice is evaluated at the transcript reached by that continuation, including any earlier interventions. A continuation may use fewer than K interventions. Once K interventions have been made, the remaining queries follow the reference. The search compares terminal regret among candidate continuations whose terminal raw balance violation is at most 10^{-8} above that of the reference trajectory from the same initial transcript and budget. This reference trajectory is fixed throughout the joint search. Terminal regret is evaluated at the final projected aggregate state using the private evaluator in Appendix A.6. The exact joint search returns the smallest terminal regret among these candidate continuations. We denote it by $R_K^{\mathrm{glob},\star}$, as in Appendix G.5.

Numerical comparisons. Ties in terminal regret are resolved deterministically using raw balance violation, local work, communication, decision position, party, and local-work level, in that order. We verify the reported terminal values by recomputing each selected trajectory.

Table 15: Exact joint-search results over 480 runs. Regret reductions are percentages. Reduction beyond the best single intervention is defined in Eq. (159). The final column is the largest reference-relative raw balance violation difference among the runs at the indicated quota.

K	Reduction vs. reference	Reduction beyond single intervention	Max ΔV_{raw}
1	16.45	0.00	-1.95×10^{-6}
2	23.59	8.55	-1.95×10^{-6}
3	26.99	12.61	-1.95×10^{-6}

Table 16: Exact joint-search results by budget. All regret reductions are percentages. The last column gives the largest reference-relative raw balance violation difference across runs and quotas $K \in \{1, 2, 3\}$ within each budget.

	Reduction vs. reference			Beyond single		
Budget	$K = 1$	$K = 2$	$K = 3$	$K = 2$	$K = 3$	Max ΔV_{raw}
Communication 35	4.72	6.34	6.92	1.70	2.31	-1.95×10^{-6}
Communication 51	8.08	17.97	19.98	10.77	12.95	-9.38×10^{-5}
Communication 75	20.21	32.25	38.58	15.08	23.02	-6.14×10^{-5}
Communication 107	35.44	49.44	58.21	21.68	35.28	-6.02×10^{-5}
Local work 864	30.39	36.33	40.12	8.53	13.97	-8.62×10^{-4}
Local work 1296	22.98	36.79	45.16	17.93	28.79	-8.32×10^{-5}

Evaluation set and comparison measures. The joint-search evaluation uses eight days separate from the 24 days used for the sequential evaluation. With five sets of resource parameters, two conflict levels, and the six budgets that permit post-initialization decisions under the reference, this gives $8 \times 5 \times 2 \times 6 = 480$ runs. The date partition is described in Appendix G.6. Tables 15 and 16 report regret reductions relative to the reference and to the best single intervention. For an evaluation set $\mathcal{U}$, the latter comparison is

$$\frac{\sum_{u \in \mathcal{U}} \left(R_{1,u}^{\mathrm{glob},\star} - R_{K,u}^{\mathrm{glob},\star} \right)}{\sum_{u \in \mathcal{U}} R_{1,u}^{\mathrm{glob},\star}}, \tag{159}$$

when the denominator is positive. Here u indexes runs. This ratio measures the relative reduction in aggregate regret from the best single intervention to the joint solution with quota K. Relative to the reference, aggregate regret falls by 16.45%, 23.59%, and 26.99% at $K = 1, 2, 3$, respectively. Relative to the best single intervention, the reductions at $K = 2$ and $K = 3$ are 8.55% and 12.61%. Both quotas improve strictly on the best single intervention in 478 of the 480 runs. These are the joint-search results summarized in Figure 3b. Aggregate regret decreases with the intervention quota

in every budget setting in Table 16. The largest raw balance violation difference is negative in every setting, so all selected trajectories have lower terminal raw balance violation than their references. Multiple interventions therefore remain beneficial when selected jointly.

I Reference-Anchored Query Selection

This appendix describes the reference-anchoring experiment introduced in Section 6. At each decision, a calibrated rule selects either a proposed adaptive query or the next reference query. We report the rate of objective deterioration relative to the reference, the distribution of terminal regret, and how often adaptive queries are accepted. The experiment uses the separate coordination setting described in Appendix G. Reference-relative policy improvement also motivates safe multi-agent and offline learning methods (Bianchi et al., 2024; Wei et al., 2024). Here Learn-then-Test calibrates the query-selection rule against a day-level objective-deterioration target (Angelopoulos et al., 2025).

I.1 Selecting Queries Relative to the Reference

Queries and resource limits. The experiment has eight parties, with resource counts and query costs listed in Table 6. A query is $q = (m, J)$, where m identifies the party and J is its local-work level. For party m with R_m resources, the cost parameters in Eq. (139) are $c_m^{\mathrm{comm}} = 1$ and $c_m^{\mathrm{loc}} = R_m$, giving one communication unit and JR_m local-work units per query. Every selected query must satisfy both resource limits in Eq. (140). Each trajectory starts with one query to every party at $J = 2$. This initialization costs eight communication units and 64 local-work units. The responses enter the transcript, and initialization counts toward the total resource budget. After response Y_t, the coordination action A_t selects the next query q_{t+1}, following Appendix A.2. This choice uses the budget remaining after q_t has been charged.

Adaptive proposals and the reference choice. At the current post-response transcript τ_t^+, write q_{t+1}^{ad} for the adaptive proposal and q_{t+1}^{ref} for the reference query. The adaptive proposal and its risk score are computed from the public coordination state and resource information in this transcript. The reference follows the static schedule selected on the reference-validation data, as described in Appendix G.4. The anchored rule compares the score with its calibrated acceptance threshold. It selects the proposed query when accepted and the reference query otherwise. After receiving the selected query's response, the coordinator updates its state and applies the rule at the next decision. Adaptive and reference queries can therefore occur within the same trajectory. The threshold remains fixed throughout test evaluation. The reference query is evaluated at the transcript reached by the anchored policy, including previous accepted queries and their resource use. Selecting a different query counts as an intervention under Eq. (26). The query cost is charged when the query is issued.

Paired terminal comparisons. We compare three policies on each scenario: the static reference π_{ref}, the adaptive policy without the acceptance rule π^{ad}, and the anchored policy π^{anc}. Each comparison uses the same private instance, initialization, resource budget, public aggregate projection, and private-evaluator scoring rule. The terminal reference value comes from a separate run of π_{ref} from this common initialization. It remains the comparison value throughout that scenario, while the reference query offered at each intermediate decision is evaluated at the current transcript.

I.2 Objective Deterioration and Risk Calibration

Scenario and daily losses. A scenario w specifies a calendar day, a set of resource parameters, a conflict level, and a budget setting. Following Appendix G.5, let $F_{\mathrm{raw}}(\pi; w)$ denote the normalized private-evaluator objective at the terminal cached responses before public aggregate projection under policy π. Objective values are computed offline from the private instance, using the same positive normalization factor for all policies within a scenario. For policy π, the deterioration indicator in Eq. (155) is

$$U_w(\pi) = \mathbf{1}\left\{F_{\mathrm{raw}}(\pi; w) > F_{\mathrm{raw}}(\pi_{\mathrm{ref}}; w)\right\}. \tag{160}$$

This indicator records whether the private-evaluator objective at the cached terminal responses increases relative to the corresponding reference value before public projection, irrespective of the size of the increase. The reported test deterioration rates use $\pi = \pi^{\mathrm{anc}}$. Let j index calendar days and

let $\mathcal{W}_j$ contain the scenarios for day j. Each day has five sets of resource parameters, three conflict levels, and ten budget settings, so $|\mathcal{W}_j| = 150$. The daily loss is the fraction of these scenarios with objective deterioration:

$$L_j(\pi) = \frac{1}{150} \sum_{w \in \mathcal{W}_j} U_w(\pi) \in [0, 1]. \tag{161}$$

For a set $\mathcal{D}$ of day indices, the empirical deterioration rate is

$$\widehat{\mathcal{R}}_{\mathrm{anc},\mathcal{D}}(\pi) = \frac{1}{|\mathcal{D}|} \sum_{j \in \mathcal{D}} L_j(\pi) = \frac{1}{150|\mathcal{D}|} \sum_{j \in \mathcal{D}} \sum_{w \in \mathcal{W}_j} U_w(\pi). \tag{162}$$

Because every day contributes 150 scenarios, this average equals the rate obtained by pooling all scenario indicators. Statistical inference treats each day, represented by $L_j(\pi)$, as one observation.

Calibration and its risk target. The chronological intervals are given in Table 9. Risk calibration uses 183 complete days from 2019-07-01 to 2020-01-01. The selected rule is evaluated on the 272 complete test days from 2020-01-01 to 2020-10-01. Both intervals exclude the end date. Reference selection and policy development use the earlier intervals in that table. Calibration targets the expected daily loss,

$$\mathcal{R}_{\mathrm{anc}}(\pi) = \mathbb{E}[L_j(\pi)], \tag{163}$$

where the expectation is over a new day, with the five parameter sets, three conflict levels, and ten budget settings held fixed. Each combination in this experimental grid receives equal weight. The target is $\alpha = 0.05$. In the Learn-then-Test (LTT) procedure (Angelopoulos et al., 2025), let γ index candidate threshold settings and let π_γ denote the policy induced by each setting. Each candidate policy is evaluated using its daily losses $L_j(\pi_\gamma)$ over the calibration days. LTT tests $H_{0,\gamma} : \mathcal{R}_{\mathrm{anc}}(\pi_\gamma) > \alpha$. With valid p-values and a multiple-testing procedure that controls the family-wise error probability at level δ_{cal}, the probability, over the calibration sample, of retaining any setting with risk above α is at most δ_{cal}. Here α is the target for the expected daily loss, and δ_{cal} controls the probability of retaining a setting that exceeds this target. Let $\widehat{\gamma}$ denote the setting selected from those retained by LTT. The anchored policy $\pi^{\mathrm{anc}} = \pi_{\widehat{\gamma}}$ uses this fixed setting throughout test evaluation. The standard finite-sample interpretation assumes independent days drawn from a common distribution. This assumption also underlies the test p-value reported below. Scenarios within a day may be dependent, as they enter the analysis through the single bounded loss $L_j(\pi)$.

Relation to the other evaluation measures. The deterioration event in Eq. (160) concerns the objective before public aggregate projection. Raw balance violation in Eq. (146) measures the discrepancy from the shared constraint at the cached terminal aggregate responses. Terminal regret R_π in Eq. (145) uses the projected aggregate state. We report its empirical 95th percentile, denoted P95, alongside the deterioration rate. The LTT risk target applies to the expected daily deterioration loss in Eq. (163). P95 summarizes the upper tail of the observed terminal regret distribution.

I.3 Test Evaluation and Results

Evaluation set. The test combines 272 days, five sets of resource parameters, three conflict levels $\{0.5, 0.75, 1.0\}$, and ten budget settings, giving 40,800 scenarios. Communication budgets are $\{8, 12, 16, 24, 32\}$ and local-work budgets are $\{64, 128, 256, 512, 1024\}$. The communication and local-work limits are varied separately, with the other limit held fixed as specified in Appendix G.3. Within each budget setting, the three policies are evaluated on the same 4,080 scenarios. Confidence intervals for paired mean regret differences use the hierarchical bootstrap described in Appendix G.6, with 50,000 replicates obtained by resampling parameter sets first and then contiguous seven-day blocks. The risk test uses the daily losses in Eq. (161).

Deterioration rate and terminal regret. Across all 40,800 test scenarios, the anchored policy's cached-response objective before public projection exceeds the corresponding reference value in 1.93% of scenarios. The day-level test of the null hypothesis $\mathcal{R}_{\mathrm{anc}}(\pi^{\mathrm{anc}}) > 0.05$ gives $p = 0.0303$. Table 17 reports the empirical P95 regret of the three policies within each budget setting. Each percentile uses the 4,080 terminal regrets in that setting, pooling days, parameter sets, and conflict levels. The deterioration column averages $U_w(\pi^{\mathrm{anc}})$ over the same scenarios. The final column

Table 17: Test results by budget. P95 is the empirical 95th percentile of normalized terminal regret, reported as a percentage. Deterioration measures cached-response objective increases before public projection. The final column reports trajectories with at least one accepted adaptive query.

Budget	Reference P95	Adaptive P95	Anchored P95	Deterioration (%)	Trajectories with acceptance (%)
Communication 8	1.808	1.808	1.808	0.00	0.00
Communication 12	1.808	2.036	1.809	0.05	0.17
Communication 16	1.924	1.925	1.881	13.33	29.31
Communication 24	0.915	1.625	0.915	1.00	4.80
Communication 32	0.534	1.713	0.534	2.57	5.00
Local work 64	1.808	1.808	1.808	0.00	0.00
Local work 128	1.808	2.034	1.808	0.00	0.00
Local work 256	1.812	2.207	1.811	2.30	6.69
Local work 512	0.741	2.224	0.741	0.00	0.00
Local work 1024	0.534	1.796	0.534	0.00	0.00

gives the percentage of trajectories containing at least one accepted adaptive query. At the reported precision, the anchored P95 equals the reference P95 in seven budget settings. It is lower by 0.043 percentage points at communication budget 16 and by 0.001 percentage points at local-work budget 256. At communication budget 12, it is higher by 0.001 percentage points. Compared with the adaptive policy without anchoring, the anchored policy has lower empirical P95 regret in eight budget settings and equal P95 in the other two. When this comparison is made separately within each conflict level, anchoring gives lower P95 in 23 of the 30 combinations of conflict level and budget: eight at 0.5, seven at 0.75, and eight at 1.0. Deterioration rates vary across budget settings. At communication budget 16, the rate is 13.33%, compared with 1.93% across all ten budgets. The calibration target $\alpha = 0.05$ applies to the expected daily loss, which gives equal weight to all ten budget settings.

I.4 Frequency of Accepted Adaptive Queries

At communication budget 16, 29.31% of test trajectories contain at least one accepted adaptive query, the highest rate in Table 17. No adaptive query is accepted at five of the ten budget settings. Two of these, communication budget 8 and local-work budget 64, are exhausted by the fixed initialization. On 83.8% of test days, at least one of the 150 scenarios contains an accepted adaptive query. Across all test trajectories of the anchored policy, 1.32% of post-initialization query choices differ from the reference choice at the current transcript.

J Transcript Rules and Transfer across Resource Budgets

This appendix gives the methods and detailed results for the timing and transfer studies summarized in Section 6.3. The experiments examine whether transcript information helps identify when to replace a reference query and how well predictors generalize to budgets excluded from training. In the trigger-rule experiments, an instance-aware evaluator supplies an alternative query at each decision, and the rule chooses between this query and the reference query. Separate prediction experiments compare learned scores on fixed sets of recorded decision points. We use the main coordination setting in Appendix G. Under the reference schedule, six budget settings allow decisions after initialization: communication budgets $35, 51, 75, 107$ and local-work budgets $864, 1296$. The resource parameters, conflict levels, initialization, and query costs are specified in Appendix G.2 and Appendix G.3. Terminal regret is measured at the projected aggregate state, while raw balance violation is measured from the cached responses before projection, as in Appendix G.5.

J.1 Candidate Queries and Continuations

Reference and replacement queries. At the post-response transcript τ_t^+, the reference action selects q_{t+1}^{ref}. A query has the form $q = (m, J)$, where $m \in \{1, \ldots, M\}$, $M = 8$, and $J \in \{2, 5, 10, 25, 50\}$ is the local-work level. Its communication and local-work costs are c_m^{comm} and

Jc_m^{loc}, respectively. The set of affordable alternatives is

$$\mathcal{Q}_{\mathrm{alt}}(\tau_t^+) = \left\{ (m, J) : \begin{array}{l} m \in \{1, \dots, M\}, \quad J \in \{2, 5, 10, 25, 50\}, \\ (m, J) \neq q_{t+1}^{\mathrm{ref}}, \\ (m, J) \text{ satisfies Eq. (140)} \end{array} \right\}. \tag{164}$$

Affordability uses the budget remaining after q_t has been charged. This set is determined by the public query costs, remaining budget, and reference query. Selecting a query different from the reference counts as an intervention. Its resource cost is charged when it is issued, following Appendix A.2.

The evaluator-selected alternative. For a post-response transcript τ, let $V_{\mathrm{raw}}^{\mathrm{ref}}(\tau)$ be the terminal raw balance violation obtained by following the reference from τ with the remaining budget. For each $q \in \mathcal{Q}_{\mathrm{alt}}(\tau)$, the evaluator computes the response to q, updates the coordination state, and then follows the reference. Denote the terminal regret and raw balance violation of this continuation by $\overline{R}(\tau; q)$ and $\overline{V}_{\mathrm{raw}}(\tau; q)$. The evaluator retains the queries in $\mathcal{Q}_{\mathrm{ret}}(\tau) = \left\{ q \in \mathcal{Q}_{\mathrm{alt}}(\tau) : \overline{V}_{\mathrm{raw}}(\tau; q) \leq V_{\mathrm{raw}}^{\mathrm{ref}}(\tau) + 10^{-8} \right\}$. Among these queries, it selects the one with the smallest terminal regret. Regret ties are resolved by minimizing, in order, terminal raw balance violation, total local-work cost, total communication cost, party index, and local-work level. Write $q^{\mathrm{alt}}(\tau)$ for the selected query. If no alternative is retained, the reference continues. The evaluator uses the private instance to compute these terminal values and select $q^{\mathrm{alt}}(\tau)$. In the trigger-rule experiments, the rule decides whether to execute this query or the reference query.

Continuations used to evaluate a choice. A decision record v specifies a post-response transcript τ_v and a remaining intervention quota k_v before the current choice. Write $q^{\mathrm{ref}}(\tau_v)$ for the reference query at this transcript. Keeping the reference executes $q^{\mathrm{ref}}(\tau_v)$ and leaves k_v interventions available. Replacing it is allowed when $k_v > 0$ and executes $q^{\mathrm{alt}}(\tau_v)$, leaving $k_v - 1$ interventions available. In both cases, the response is computed at the current coordination state, which is then updated. The three continuation rules specify how the trajectory proceeds after this response. A reference continuation follows the selected static schedule from the resulting transcript and remaining budget. An exact continuation recursively considers the reference query and the evaluator-selected alternative at each subsequent decision. The evaluator recomputes the alternative at each reached transcript. The search uses at most the quota remaining after the current choice and follows the reference once that quota is exhausted. It selects the smallest terminal regret among continuations whose terminal raw balance violation is at most $V_{\mathrm{raw}}^{\mathrm{ref}}(\tau_v) + 10^{-8}$. The comparison value $V_{\mathrm{raw}}^{\mathrm{ref}}(\tau_v)$ remains fixed throughout the search. A fixed-rule continuation applies the same specified trigger rule at subsequent decisions after either current choice. The exact continuation in this appendix searches the reference query and one evaluator-selected alternative at each state. The joint search in Appendix H.4 considers all affordable alternative queries.

Table 18: Continuations and evaluations in the transcript-learning experiments. Each branch begins at the same decision record. An intervention changes the state from which later choices are made.

Comparison	Continuation used for values	Evaluation
Position ranking and residual rule	The current query is followed by the reference.	Position ranking uses fixed trajectory records. The trigger rule is also executed along complete trajectories, with a new alternative supplied at each reached state.
Absolute continuation regret	The first query is fixed; later reference and replacement choices are searched exactly with the remaining quota.	The fitted predictor chooses between the two current queries at each reached state.
Direct advantage prediction and representation comparisons	Both branch values use the same exact continuation search.	Scores select records from a fixed evaluation collection.
Branch values under a fixed trigger rule	Both current choices are followed by the same fixed residual-based rule.	A separate greedy rule is executed using these branch values.

J.2 Advantages and Evaluation Measures

Advantage of replacing a query. For decision record v and remaining quota k_v, let $R_{\mathrm{keep}}(v, k_v)$ and $R_{\mathrm{trigger}}(v, k_v)$ denote the terminal regrets after keeping or replacing the reference query, respectively. Both use the continuation specified for the experiment in Table 18. The replacement advantage is

$$A_v = R_{\mathrm{keep}}(v, k_v) - R_{\mathrm{trigger}}(v, k_v), \tag{165}$$

as in Eq. (152). A positive value means that replacing the query reduces terminal regret under that continuation. With a reference continuation, subsequent queries follow the reference schedule. With an exact continuation, later interventions are allowed within the quota remaining after the current choice.

Net capture on fixed decision records. Let $\mathcal{D}$ be an evaluation set of $N = |\mathcal{D}|$ decision records, and let $I \subseteq \mathcal{D}$ contain the records selected by a score or threshold. Following Eq. (153), net capture is $\mathrm{NC}(I) = \frac{\sum_{v \in I} A_v}{\sum_{v \in \mathcal{D}} [A_v]_+}$, $[a]_+ = \max\{a, 0\}$, when the denominator is positive. All scores compared on $\mathcal{D}$ use the same advantages and denominator. Selecting a record with negative advantage reduces the numerator. The direct-advantage results report the contributions from beneficial and harmful selections separately:

$$\mathrm{PosCapture}(I) = \frac{\sum_{v \in I} [A_v]_+}{\sum_{v \in \mathcal{D}} [A_v]_+}, \qquad \mathrm{FalseCost}(I) = \frac{\sum_{v \in I} [-A_v]_+}{\sum_{v \in \mathcal{D}} [A_v]_+}. \tag{166}$$

Thus $\mathrm{NC}(I) = \mathrm{PosCapture}(I) - \mathrm{FalseCost}(I)$. Trigger recall is the fraction of records with $A_v > 0$ that are selected. Keep recall is the fraction with $A_v \le 0$ that are not selected. Recall counts records, while the capture measures account for the magnitudes of their advantages.

Ranking net capture. For a score g and selection rate q, let $\mathrm{Top}_q(g)$ contain the $\mathrm{round}(qN)$ highest-scoring records, using a fixed order to resolve ties. Following Eq. (154), ranking net capture averages net capture over five selection rates:

$$\mathrm{RNC}(g) = \frac{1}{5} \sum_{q \in \{0.02, 0.05, 0.10, 0.20, 0.30\}} \mathrm{NC}\big(\mathrm{Top}_q(g)\big). \tag{167}$$

Figure 4 uses this measure to compare the rankings produced by different predictors.

Capture from a fixed number of ranked positions. For the position-ranking comparison in Section 6.3, let u index the evaluated reference trajectories and let $\mathcal{P}_u$ contain the evaluated positions in trajectory u. Let $T_{u,K}$ contain the $\min\{K, |\mathcal{P}_u|\}$ highest-scoring positions in that trajectory. Position-ranking capture is

$$\mathrm{RankCapture}_K = \frac{\sum_u \sum_{v \in T_{u,K}} [A_v]_+}{\sum_u \sum_{v \in \mathcal{P}_u} [A_v]_+}, \tag{168}$$

when the denominator is positive. The advantages here use reference continuations. This measure gives the fraction of total positive advantage contained in the selected positions. Selected positions with nonpositive advantage contribute zero. The 47.10% result uses $K = 3$.

Capture by an executed policy. Policy capture evaluates a rule through its terminal regret after execution. At each reached transcript, the evaluator supplies an alternative query and the rule decides whether to use it. For an evaluation set $\mathcal{U}$ of runs, let $R_{\pi,u}$ be the terminal regret of policy π in run u, and let $R_{\mathrm{ref},u}$ be the terminal regret of the reference. The policy may use at most K interventions. The benchmark $R^{\mathrm{eval}}_{K,u}$ is the minimum terminal regret returned by the exact evaluator with at most K interventions in the same run. The evaluator uses the candidate choices specified for that comparison, with the same private instance, initialization, and resource budget. Following Eq. (149) and the aggregation convention in Appendix G.5, aggregate policy capture is

$$\mathrm{PolicyCapture}_K(\pi; \mathcal{U}) = \frac{\sum_{u \in \mathcal{U}} (R_{\mathrm{ref},u} - R_{\pi,u})}{\sum_{u \in \mathcal{U}} (R_{\mathrm{ref},u} - R^{\mathrm{eval}}_{K,u})}, \tag{169}$$

when the denominator is positive. The comparison beyond the best single intervention uses $\mathrm{MultiCapture}$ in Eq. (150), with benchmark values computed for the evaluated runs. Trigger rate is the number of executed interventions divided by the number of eligible positions reached by the rule. The beneficial-trigger rate is the fraction of those interventions with positive advantage under the continuation specified for the experiment.

J.3 Data Partitions and Prediction Models

The fixed-record prediction studies use days from the policy-development interval in Appendix G.6. The online timing evaluation uses the later dates listed below. For model selection, all records from a given day remain in the same fold. The reference-anchoring test in Appendix I is a separate experiment and is not used for model selection here.

Within-budget and held-out-budget comparisons. A within-budget model is fitted and evaluated on the same budget using separate days. For leave-one-budget-out evaluation, abbreviated LOBO, one of the six budgets is excluded from model fitting, model selection, and threshold selection. The other five budgets provide the development data, and the held-out budget is used for evaluation. The predictors use budget and progress as inputs. The representation comparisons use 12 fitting days and six days for model selection. The matched-selection-rate comparison uses three-fold out-of-fold predictions over all 18 fitting and model-selection days to choose the common selection rate. Its evaluation records come from six separate days. The direct-advantage study uses 82,213 training pairs and 27,690 evaluation pairs. The matched-rate and representation comparisons use the same evaluation pairs and exact branch advantages.

Calendar dates for prediction studies. The fitting dates are `2018-01-19, 01-25, 01-30, 02-08, 03-22, 04-24, 04-29, 05-19, 06-07, 06-09, 07-14, 08-11`; model-selection dates are `2018-09-18, 09-21, 10-06, 10-17, 11-11, 12-21`; and evaluation dates are `2018-02-05, 09-12, 10-12, 12-02, 12-05, 12-18`. The online timing comparison in Table 19 uses 48 later dates: January `01,11,13,17,18,23,25,27,29`; February `04,06,14,20`; March `03,05,06,08,10,16,17,25,28`; April `07,11,14,18,19,20,27`; May `09,14,15,16,19,21,24,27,28`; and June `02,04,09,13,14,16,19,24,26,29`, all in 2019.

Prediction models. For direct advantage prediction, we select among ridge regression with ℓ_2 penalty 1, histogram gradient boosting (HGB), and a value-weighted version of the same HGB model. HGB uses learning rate 0.05, 120 boosting iterations, at most 15 leaves, a minimum of 64 samples per leaf, ℓ_2 regularization 1, and no early stopping. Value weights are $|A|$ divided by the median nonzero $|A|$, clipped to $[0.25, 8]$ and normalized to mean one. Model selection uses three folds grouped by day within the five development budgets. Unweighted HGB is selected for each held-out budget. The feed-forward screen uses standardized inputs and ReLU multilayer perceptrons with hidden widths (64) or $(128, 64)$, Adam with learning rate 10^{-3}, ℓ_2 penalty 10^{-4}, at most 100 epochs, and early stopping with a 10% validation fraction and patience 12. For T2, a linear 64-dimensional static encoder is concatenated with the last state of a one-layer GRU with hidden dimension 32 applied causally to 172-dimensional public-history tokens. A 64-unit head predicts advantage. This model is trained for 12 epochs with batch size 512, AdamW at learning rate 10^{-3} and weight decay 10^{-5}, Smooth-ℓ_1 loss with transition point 0.5, and gradient norm clipped at 1. T3 uses the same HGB configuration as the direct-advantage models. The prediction target is branch regret or trigger advantage, depending on the experiment. All preprocessing statistics are fitted only on the corresponding training folds.

J.4 Residual-Based Trigger Rules

Public residual and its normalized change. Let $\widetilde{X}_t = (\widetilde{x}^{(t)}_{mh})_{m,h} \in \mathbb{R}^{M\times 24}$ collect the most recent hourly responses from all parties after query q_t, using the notation of Appendix G.3. Let $\mathbf{Z}_t \in \mathbb{R}^{M\times 24}$ be their coordinated projection in the public state. Define the residual matrix $r_t = \widetilde{X}_t - \mathbf{Z}_t$. Both matrices are available to the coordinator. Terminal raw balance violation measures the aggregate hourly imbalance at termination, as defined in Eq. (146). For the public demand vector $d = (d_h)_{h=1}^{24}$,

define

$$c_d = \max\left\{\frac{1}{24}\sum_{h=1}^{24}|d_h|, 10^{-8}\right\}, \qquad \delta_t = \left(\frac{1}{24M}\left\|\frac{r_t - r_{t-1}}{c_d}\right\|_F^2\right)^{1/2}. \tag{170}$$

Here $\|\cdot\|_F$ is the Frobenius norm. The statistic δ_t is the root mean square change in the residual across parties and hours, normalized by the demand scale c_d. When no preceding residual is available, the residual difference is set to zero.

Initial values and running score. Initial location and scale are fitted separately for the two budget types: communication and local work. Let D_δ be the corresponding development sample of normalized changes, and let $D_\delta^+ = \{\delta \in D_\delta : \delta > 0\}$. Define

$$\mu_0 = \text{median}(D_\delta), \tag{171}$$

$$s_{\min} = \begin{cases}\max\{\text{Quantile}_{0.25}(D_\delta^+), 10^{-8}\}, & D_\delta^+ \neq \varnothing,\\ 10^{-8}, & D_\delta^+ = \varnothing,\end{cases} \tag{172}$$

$$s_0 = \max\left\{\text{median}_{\delta\in D_\delta}|\delta - \mu_0|, s_{\min}\right\}. \tag{173}$$

The quantile is the empirical 25th percentile of the positive changes. When an estimate for one budget type is unavailable, the pooled development estimate is used. The rule uses $\beta = 0.9$ and computes the score before updating the running location and scale:

$$e_t = \text{clip}_{[-12,12]}\left(\frac{\delta_t - \mu_{t-1}}{\max\{s_{t-1}, s_{\min}\}}\right), \tag{174}$$

$$\mu_t = \beta\mu_{t-1} + (1-\beta)\delta_t, \tag{175}$$

$$s_t = \beta s_{t-1} + (1-\beta)|\delta_t - \mu_{t-1}|. \tag{176}$$

The score measures how far the current residual change lies above or below the preceding running location, relative to the running scale. Clipping restricts the score to $[-12, 12]$. The fitted initial values and β remain fixed during evaluation, while μ_t and s_t are updated as responses arrive within each trajectory.

Progress and percentile calibration. Let H_0 be the initial reference horizon used by the rule, and let n_t index the scored positions. Reference progress is $p_t = \text{clip}_{[0,1]}\left(\frac{n_t}{\max\{H_0-1,1\}}\right)$. The three progress phases are $[0, 1/3)$, $[1/3, 2/3)$, and $[2/3, 1]$. The rule maps e_t to a percentile $u_t \in [0, 1]$ using the score distribution estimated from development records. The estimate uses 201 evenly spaced quantile levels. A separate distribution is estimated for each combination of budget type and progress phase when that group contains at least 32 records. For smaller groups, the rule uses the distribution for the corresponding budget type. If that estimate is unavailable, it uses the pooled distribution. These distributions remain fixed during evaluation.

Trigger threshold. The rule sets $K_0 = \min\{3, \lfloor 0.1H_0\rfloor\}$, $\theta_t = \max\left\{0.8,\ 1 - \frac{K_0 - a_t}{\max\{H_0 - n_t, 1\}}\right\}$, where a_t counts preceding trigger attempts. When a retained alternative is available and intervention quota remains, the rule triggers if $u_t \geq \theta_t$. The percentile threshold is at least 0.8. The executed rule makes at most three interventions and follows the reference once its intervention quota is exhausted.

Phase-matched random rule. The random rule uses the residual rule's trigger-attempt frequencies on development data, estimated separately for each budget type and progress phase. The probabilities in phase order are

$$\begin{aligned}\text{Communication:}&\quad (0.1015625,\ 0.0781250,\ 0.0124975),\\ \text{Local work:}&\quad (0.0550000,\ 0.0774554,\ 0.0211169).\end{aligned} \tag{177}$$

At each eligible decision, the rule makes an independent Bernoulli draw with the probability for its budget type and progress phase. When it triggers, it executes the evaluator-selected alternative. The rule makes at most three interventions. Each run uses a random seed derived from the fixed base seed 99173. The observed trigger rates during evaluation are 5.31% for the random rule and 5.99% for the residual rule.

Results and other baselines. Table 19 compares the residual rule, raw dual change, phase-matched random triggering, and a predictor using budget and progress. It also reports the position-ranking result. All executed rules use evaluator-selected alternatives and at most three interventions. In the

Table 19: Transcript rules with evaluator-selected alternatives. All numerical entries are percentages. Position-ranking capture uses Eq. (168). Policy capture evaluates complete trajectories. Beneficial triggers are executed interventions with positive advantage. A dash indicates that the position-ranking result is not reported for that rule.

Rule	Position-ranking capture at 3	Policy capture at 3	Trigger rate	Beneficial triggers
Normalized residual change	47.10	26.46	5.99	80.02
Raw dual change	—	25.89	10.14	71.85
Phase-matched random	—	11.56	5.31	73.46
Budget and progress	—	6.04	1.21	66.94

offline ranking diagnostic, the three highest-scoring positions in each reference trajectory contain 47.10% of the total positive advantage. Position-ranking capture ranges from 46.31% to 47.53% across resource-parameter sets and from 46.66% to 47.57% across conflict levels. The online percentile-threshold rule achieves 26.46% policy capture with a trigger rate of 5.99%. Raw dual change achieves 25.89% policy capture with a trigger rate of 10.14%. The residual rule's trigger rate is about 41% lower. The comparison contains 2,880 trajectories sharing 48 calendar dates. A paired nonparametric bootstrap resamples dates and moves all seed, conflict, and budget trajectories for a sampled date together. With 20,000 replicates and fixed seed 20271021, the 95% intervals are $[22.36, 30.59]\%$ for residual-rule capture and $[18.54, 32.78]\%$ for raw-dual capture. The paired capture difference is 0.57 percentage points with interval $[-7.17, 8.75]$. It is therefore not interpreted as a capture advantage. The residual-minus-raw-dual trigger-rate difference is -4.15 points with interval $[-4.76, -3.53]$. Its policy capture is lower than that of raw dual change for one resource-parameter set and at the higher conflict level. Using target-budget score quantiles for recalibration gives aggregate policy capture of 19.67%, compared with 26.46% under the original calibration.

J.5 Predicting Branch Regret and Trigger Advantage

Table 20: Executed policy based on predicted continuation regret, with at most three interventions. Both capture measures are percentages and are reported without clipping. Multi-intervention capture uses Eq. (150).

Held-out budget	Policy capture at 3	Multi-intervention capture
Communication 35	-57.12	-466.12
Communication 51	23.54	-32.47
Communication 75	28.52	-54.11
Communication 107	38.65	-88.75
Local work 864	26.10	-169.43
Local work 1296	-9.14	-141.49
Aggregate	8.11	-144.96

Absolute continuation regret. The first prediction target is the terminal regret of each current choice under the exact continuation in Table 18. For the current decision record v, the predictor produces $\widehat{R}_{\text{keep}}(v, k_v)$ and $\widehat{R}_{\text{trigger}}(v, k_v)$. The rule selects the query with the smaller predicted terminal regret, subject to the remaining intervention quota. After receiving the response, it makes a new prediction from the updated transcript. The aggregate policy capture is 8.11%, and the multi-intervention capture is -144.96%. The policy reduces aggregate regret relative to the reference but has higher aggregate regret than the best single intervention. The comparison with the best single intervention is also unfavorable in every reported resource-parameter and conflict-level aggregate.

Direct prediction of trigger advantage. The second target is A_v in Eq. (165), computed using the same exact continuation rule for both current choices. A predictor is fitted to this difference, and

Table 21: Direct advantage prediction on held-out budgets. Capture, false cost, and recall entries are percentages. Positive capture and false cost are defined in Eq. (166). Spearman is the rank correlation between the predicted score and exact advantage.

Held-out budget	Positive capture	Net capture	False cost	Trigger recall	Keep recall	Spearman
Communication 35	21.8	−23.4	45.2	26.8	91.1	0.288
Communication 51	2.0	1.7	0.3	7.2	99.4	0.373
Communication 75	22.3	11.1	11.2	22.4	96.9	0.377
Communication 107	51.7	30.6	21.0	39.2	93.0	0.509
Local work 864	12.2	5.1	7.1	6.4	98.4	0.360
Local work 1296	27.1	−3.6	30.7	21.6	95.8	0.313

a threshold selects decision records on the held-out budget. Table 21 reports the capture and recall of these selections on the fixed evaluation records. Spearman rank correlation is positive in all six budgets, ranging from 0.288 to 0.509. Net capture ranges from -23.4% to 30.6%. At communication budget 35, harmful selections outweigh the captured positive advantage. At communication budget 51, false cost is 0.3%, and the selected records capture 2.0% of the available positive advantage.

Branch values under a fixed continuation rule. A further comparison evaluates keeping and replacing the current query with the same residual-based trigger rule used for subsequent decisions after either choice. The evaluator computes both branch values directly from continuations under this fixed rule. Among 4,267 evaluated eligible states, 58.80% have positive replacement advantage under this continuation. A greedy rule uses these branch values to choose a query at each reached state. It reduces regret relative to the reference by 4.01%, compared with 4.54% for the best single intervention in this comparison. The reported maximum raw balance violation increases by 0.01390.

J.6 Public and Privileged Representations

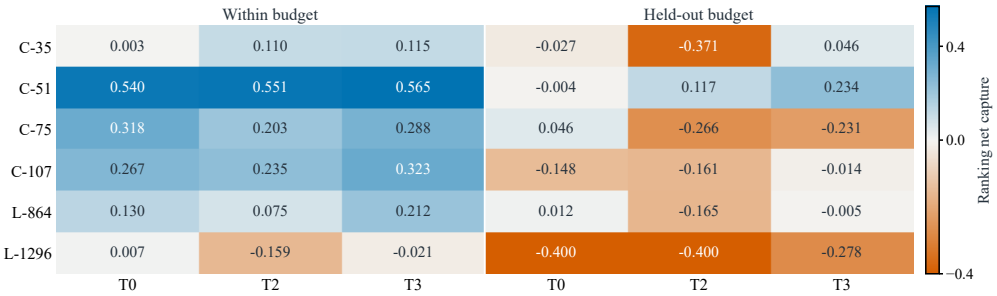


Figure 4: Ranking net capture (RNC) within and across budgets. RNC averages net capture after selecting the highest-scoring 2%, 5%, 10%, 20%, and 30% of the fixed evaluation records. T0 uses the compact transcript summary, T2 the full public transcript, and T3 evaluator-only private parameters and exact next responses in addition to the current public state. C and L denote communication and local-work budgets.

The representation comparison uses the replacement advantage A_v in Eq. (165), with both branch values computed using the exact continuation rule in Table 18. We fit predictors using three representations of public information and one representation with additional private information:

T0: compact transcript summary. Budget and progress, the reference query, recent residual and dual summaries, summaries of cached responses, and features describing differences between the compared queries.

T1: current public state. T0 together with all public variables retained in the current coordinator state, including per-party cached responses, their ages, resource use, and coordination variables.

T2: full public transcript. T1 together with the ordered history of public queries, responses, residuals, dual changes, and resource use. A causal recurrent model encodes this sequence.

T3: current state with private information. T1 together with private cost parameters and the responses that the two compared queries would produce at the current state, computed by the evaluator. These additional inputs are used for offline prediction.

Figure 4 compares T0, T2, and T3. Across the six within-budget comparisons, mean RNC is 0.011

Table 22: RNC for public representations. Within-budget models use separate fitting, model-selection, and evaluation days. LOBO models are fitted without the evaluated budget. Entries average over the five selection rates in Eq. (167).

	Within budget			LOBO	
Budget	T0	T1	T2	T0	T2
Communication 35	0.003	0.098	0.110	−0.027	−0.371
Communication 51	0.540	0.550	0.551	−0.004	0.117
Communication 75	0.318	0.255	0.203	0.046	−0.266
Communication 107	0.267	0.286	0.235	−0.148	−0.161
Local work 864	0.130	0.133	0.075	0.012	−0.165
Local work 1296	0.007	0.009	−0.159	−0.400	−0.400

Table 23: RNC with private-instance inputs. Best public is the largest within-budget RNC among T0, T1, and T2 in Table 22. The difference subtracts that value from T3's within-budget RNC.

Budget	Best public	T3 within budget	Difference	T3 LOBO
Communication 35	0.110	0.115	0.005	0.046
Communication 51	0.551	0.565	0.014	0.234
Communication 75	0.318	0.288	−0.030	−0.231
Communication 107	0.286	0.323	0.037	−0.014
Local work 864	0.133	0.212	0.079	−0.005
Local work 1296	0.009	−0.021	−0.030	−0.278

higher for T1 than for T0 and 0.042 lower for T2 than for T0. T2 improves on T0 at communication budgets 35 and 51 and has lower RNC in the other four budgets. Its LOBO RNC is negative in five of the six budgets. For each budget, the best public result is the largest within-budget RNC among T0, T1, and T2. T3 exceeds this value in four of the six budgets. The mean difference across all six budgets is 0.013. Its LOBO RNC is negative in four budgets.

J.7 Transfer at Matched Selection Rates

Choosing the common rate. For each budget b, we obtain out-of-fold scores from T0 models on all 18 fitting and model-selection days, using three-fold cross-validation with folds grouped by day. We rank these records by score and choose how many of the highest-scoring records to select by maximizing their total signed advantage. Selecting no records is also allowed. The selected count divided by the total number of these records defines q_b. The rate q_b is estimated from within-budget data and sets the common evaluation count for the within-budget and LOBO comparisons. The LOBO predictor is fitted using the other five budgets. Let N_b be the number of records for budget b in the separate evaluation set. The common selection count is $n_b = \mathrm{round}(q_b N_b)$. On this same evaluation set, each model selects its top n_b records, and net capture is computed for those selections. Across the six budgets, mean net capture falls from 0.358 within budget to 0.016 under LOBO evaluation, a mean decrease of 0.342, as reported in Section 6.3. The decrease occurs in all six budgets. Five within-budget values exceed 0.20, while one LOBO value does.

Relation to the representation comparison. Table 24 reports net capture at one selected rate q_b per budget. Figure 4 and Tables 22 and 23 report RNC, which averages net capture over five fixed rates. The comparisons also use different fitting procedures. The representation comparison uses 12 fitting days and six model-selection days. The matched-rate comparison uses out-of-fold predictions over all 18 days to estimate q_b. For communication budget 35, the within-budget T0 predictor in the representation comparison has net captures of -0.0003, -0.0055, 0.0258, 0.0109,

Table 24: Net capture at matched selection rates using T0. The rate is chosen from out-of-fold predictions over all 18 fitting and model-selection days, and values are computed on separate evaluation records. Difference is within-budget minus LOBO net capture. Selection rates are percentages. Capture entries are ratios.

Budget	Selection rate q_b	Within budget	LOBO	Difference
Communication 35	23.9	0.233	−0.342	0.575
Communication 51	20.3	0.681	0.090	0.591
Communication 75	17.0	0.491	0.026	0.465
Communication 107	6.2	0.424	0.239	0.185
Local work 864	13.9	0.223	0.015	0.208
Local work 1296	1.3	0.097	0.069	0.028
Mean	—	0.358	0.016	0.342

Table 25: Matched-resource response-protocol comparison. Entries are mean normalized regret multiplied by 10^3. Lower is better. The final column is the paired marginal-minus-aggregate difference with a 95% date-block bootstrap interval, on the same scale. Q counts contacts after the synchronous eight-party initialization.

Q	Aggregate response	Marginal value	Paired difference [95% interval]
0	12.367	14.329	1.962 $[1.495, 2.440]$
8	6.996	10.070	3.074 $[2.418, 3.723]$
16	3.715	8.234	4.518 $[3.707, 5.372]$
32	1.219	6.264	5.045 $[4.393, 5.740]$

and -0.0144 at the five fixed rates. Their mean gives an RNC of 0.003 at the reported precision. Under the matched-rate procedure, the selected rate is $q_b = 0.2389$, and net capture is 0.233.

K Matched-Resource Response-Protocol Comparison

This appendix specifies the protocol comparison in Section 6.3. It changes both the response rule and its corresponding public update while keeping the party-contact schedule, response dimension, local bisection work, and number of public projections fixed. The experiment therefore compares two complete, equally resourced response–update protocols. It does not isolate the response statistic or estimate the recoverability gap of the hard pair.

K.1 Two Response Interfaces

Aggregate optimizer response. The first protocol uses the conditional local problem in Eq. (132), its 24-dimensional aggregate response in Eq. (136), and the public cache, projection, and dual update in Eq. (137). Every party contact uses $J = 50$ fresh bisection steps. The coordination penalty is selected from $\rho \in \{0.25, 0.5, 1, 2, 4\}$. For an exact solve, let y_m be the returned aggregate optimizer and $v_m = z_m - u_m$ its proximal center. Optimality of the conditional problem gives

$$\rho(v_m - y_m) \in \partial V_m(y_m), \tag{178}$$

where V_m is the aggregate value function defined below. An exact aggregate response therefore identifies a subgradient at the returned proximal point. The 50-step numerical response need not satisfy Eq. (178) exactly.

Marginal-value response. For party m, let the value of an aggregate hourly allocation $a_m = (a_{mh})_{h=1}^{24}$ be

$$V_m(a_m) = \min_{x_m} \left\{ f_m(x_m) : 0 \le x_{mrh} \le \bar{x}_{mrh},\ \sum_r x_{mrh} = a_{mh} \text{ for every } h \right\}. \tag{179}$$

For the separable objective in Eq. (130), an equality multiplier satisfies $\sum_r x_{mrh}(\mu_{mh}) = a_{mh}$, with $x_{mrh}(\cdot)$ from Eq. (133). At $a_m = z_m^{(t)}$, each hour starts a fresh bracket

$$\mu_{mh}^{-} = \min_r\{c_{mrh}(p_{mrh} - \bar{x}_{mrh}) - \omega_{mrh}\}, \qquad \mu_{mh}^{+} = \max_r\{c_{mrh}p_{mrh} + \omega_{mrh}\}. \tag{180}$$

The left endpoint yields total capacity and the right endpoint yields zero. At each of 50 iterations, set $\mu = (\mu_{mh}^{-} + \mu_{mh}^{+})/2$ and $M_{mh}(\mu) = \sum_r x_{mrh}(\mu)$. If $M_{mh}(\mu) > a_{mh}$, replace μ_{mh}^{-} by μ. Otherwise, including equality, replace μ_{mh}^{+} by μ. The returned message is the negative midpoint of the final bracket:

$$\widetilde{g}_m^{(t)} = \left(-\widehat{\mu}_{mh}^{(t)}\right)_{h=1}^{24}, \qquad \widehat{\mu}_{mh}^{(t)} = \frac{\mu_{mh}^{-} + \mu_{mh}^{+}}{2}. \tag{181}$$

If the multiplier solutions form an interval, the equality branch selects toward its lower endpoint within the declared finite bracket. In exact arithmetic the limiting value is $\inf\{\mu \in [\mu_{mh,0}^{-}, \mu_{mh,0}^{+}] : M_{mh}(\mu) \le a_{mh}\}$. The implemented message uses the 50-step midpoint, not this limit. For example, two resources with $p_r = 0.4$, $c_r = 1$, $\omega_r = 0.2$, $\bar{x}_r = 1$, and target $a = 0.8$ have multiplier interval $[-0.2, 0.2]$. The returned scalar approaches $+0.2$. The numerical checks also cover targets zero and total capacity. For an exact equality multiplier, $(-\mu_{mh}^{\star})_{h=1}^{24} \in \partial V_m(z_m^{(t)})$. Eq. (181) numerically approximates this vector relation. The coordinator caches the most recent message from each party. Let $\bar{g}_h = M^{-1}\sum_m \widetilde{g}_{mh}$, $s_h = \max_m |\widetilde{g}_{mh} - \bar{g}_h|$, and

$$\widehat{g}_{mh} = \begin{cases} (\widetilde{g}_{mh} - \bar{g}_h)/s_h, & s_h > 10^{-12}, \\ 0, & s_h \le 10^{-12}. \end{cases} \tag{182}$$

After each message, the public update is

$$z^{(t+1)} = \Pi_{\mathcal{Z}}\left(z^{(t)} - \eta\left[\mathbf{1}_M(d/M)^{\top}\right] \odot \widehat{g}^{(t)}\right), \tag{183}$$

where $d = (d_h)_{h=1}^{24}$, $\mathbf{1}_M$ is the all-ones party vector, $\odot$ is elementwise multiplication, and $\Pi_{\mathcal{Z}}$ is the same bounded-balance projection as in Eq. (137). The original step fraction is selected from $\eta \in \{0.01, 0.03, 0.1, 0.3, 0.6\}$. This is a response substitution: it does not additionally reveal the aggregate optimizer response. It is also distinct from the one-bit certificate in Section 5, which is tailored to the aliased two-instance family.

K.2 Matched Resources and Data Roles

Both protocols start from the projection of the parties' aggregate preferred allocations. They then synchronously contact all eight parties once, using the same pre-batch public state, followed by $Q \in \{0, 8, 16, 32\}$ round-robin contacts. No local warm start is used. Table 26 gives the resulting resources. Communication units differ across parties in the main task, but the two interfaces contact the same parties in the same order, so their totals agree at every depth. The 24 development dates are

Table 26: Resource accounting for each interface. Local work is the sum of the declared resource-weighted bisection iterations. Message count and message dimension are identical across interfaces.

Extra contacts Q	Total messages	Scalars/message	Local work	Projections
0	8	24	4,800	2
8	16	24	9,600	10
16	24	24	14,400	18
32	40	24	24,000	34

2017-07-01, 2017-07-04, 2017-07-15, 2017-07-19, 2017-07-29, 2017-08-04, 2017-08-06, 2017-08-07, 2017-08-09, 2017-08-13, 2017-09-02, 2017-09-09, 2017-09-12, 2017-09-26, 2017-10-18, 2017-10-25, 2017-10-30, 2017-11-07, 2017-11-09, 2017-11-24, 2017-12-01, 2017-12-07, 2017-12-22, and 2017-12-23.

For each protocol, one parameter is chosen by the lowest aggregate normalized regret over all four depths on these dates. The original grids select $\rho = 2$ and $\eta = 0.01$. Because 0.01 is the lower endpoint of the original step grid, we also ran a development-only sensitivity audit on the same 24

dates, without accessing the evaluation dates. It adds a no-update baseline and four smaller positive steps. Table 27 shows that $\eta = 0.01$ remains the best marginal-protocol step on this expanded grid. The aggregate protocol at $\rho = 2$ has a development mean of 6.497×10^{-3} under the same averaging. Evaluation then uses the following 24 temporally later dates:

Table 27: Development-only marginal-step sensitivity. The score is mean normalized regret over the four query depths, five private-objective seeds, two conflict levels, and 24 development dates, multiplied by 10^3.

η	Mean regret	η	Mean regret
0	17.145	0.01	**10.087**
0.0001	16.847	0.03	15.635
0.0003	16.336	0.10	61.746
0.001	14.907	0.30	227.116
0.003	12.343	0.60	339.240

2018-02-22, 2018-02-26, 2018-03-26, 2018-04-27, 2018-05-09, 2018-05-22, 2018-06-02, 2018-06-06, 2018-06-29, 2018-07-18, 2018-07-21, 2018-08-14, 2018-08-16, 2018-08-25, 2018-09-09, 2018-09-21, 2018-09-25, 2018-09-30, 2018-10-06, 2018-10-25, 2018-11-12, 2018-11-30, 2018-12-20, and 2018-12-30.

These evaluation dates are disjoint from the dates used for the sparse-opportunity, open-loop, timing-validation, and exact joint-search results. They are held out from this comparison's tuning. Date selection depends only on the dataset checksum, date strings, and a fixed hash salt. No reserved-test data are used. Each date is evaluated under five private-objective seeds and conflict levels $\gamma \in \{0.75, 1.0\}$, giving 240 paired runs per depth. For uncertainty, we resample the 24 calendar dates with replacement and retain all ten seed–conflict runs associated with each sampled date. Table 25 reports percentile intervals from 20,000 paired replicates with fixed seed 27041. The maximum terminal balance violation over both interfaces and all evaluation runs is 1.50×10^{-16}.

L Query-Invisible Perturbations and Decision Relevance

This appendix proves Theorem 4.2 and Corollary 4.3 for the terminal-allocation submodel. Its ingredients are convex optimality and sensitivity analysis (Boyd & Vandenberghe, 2004; Shapiro, 1985). The interface conditions identify objective changes to which these tools apply.

L.1 Arbitrary Priors and Boundary Optima

Let Θ be a standard Borel instance space and ν a probability law. Assume $F : \Theta \times \mathcal{X} \to \mathbb{R}$ is measurable in θ and continuous in x, with $\mathcal{X}$ nonempty, compact, and convex. Let every F_θ be μ-strongly convex and $|F_\theta(x)| \le M(\theta)$ for all x, where $\mathbb{E}_\nu M < \infty$. There is a measurable set Θ_0 with $\nu(\Theta_0) = 1$ on which every allowed local response map is the same. This is one full-measure set for all queries, not a separate exceptional set for each query. The no-query policy and every terminal allocation in $\mathcal{X}$ are admissible. No other public state, timing observation, or solver trace carries private information. All policy randomness is independent of θ before observations. A common-randomness induction therefore gives a terminal-output law P_X that is the same for every $\theta \in \Theta_0$, including when a policy stops adaptively.

Proof of Theorem 4.2. Dominated convergence makes $\overline{F} = \mathbb{E}_\nu F_\theta$ continuous. It is also μ-strongly convex, hence has a unique minimizer $\overline{x}$. Compactness and continuity give a unique $x^\star_\theta$ for each instance. Its value is measurable by minimizing over a countable dense subset of $\mathcal{X}$. The unique minimizer is measurable by the measurable minimum theorem. The envelope makes both optimal values integrable. Independence of the terminal output from θ and Fubini's theorem give

$$\mathbb{E}_{\theta,X} F_\theta(X) = \int_{\mathcal{X}} \overline{F}(x)\, P_X(dx) \ge \overline{F}(\overline{x}). \tag{184}$$

The no-query constant decision $\overline{x}$ attains equality. The instance-aware policy selects $x^\star_\theta$, so its value is $\mathbb{E}_\nu F_\theta(x^\star_\theta)$. For completeness, strong convexity gives quadratic growth at a constrained minimizer

without assuming differentiability. Apply strong convexity to $(1-t)x_\theta^\star + tx$, compare with the minimum value, divide by t, and let $t \downarrow 0$. The result is

$$F_\theta(x) - F_\theta(x_\theta^\star) \geq \frac{\mu}{2}\|x - x_\theta^\star\|^2. \tag{185}$$

Evaluate this at $\overline{x}$ and average. A nonnegative random variable has zero expectation exactly when it vanishes almost surely. Uniqueness then gives the asserted zero-gap criterion. □

The value identity is valid for a specified prior. It does not provide an algorithm for computing $\overline{F}$ when the prior model is unknown. The lower bound also has a prior-free form. Let an objective class $\mathfrak{F}$ contain two μ-strongly convex objectives F_+, F_- with the same response maps and the same public protocol. Any transcript policy has a common output law under these two members. Thus

$$\inf_{\pi \in \Pi_{\mathrm{tr}}} \sup_{F \in \mathfrak{F}} \mathbb{E}[F(X^\pi) - F^\star] \geq \frac{\mu}{8}\|x_+^\star - x_-^\star\|^2. \tag{186}$$

Indeed, the supremum dominates the equally weighted excess loss of the pair. Quadratic growth and the squared-distance identity below give the bound. The coordinator need not know a finite candidate list. This establishes a lower bound, not the exact minimax value of the entire class.

L.2 Setting and Common Local Responses

Let $D_m \subset \mathbb{R}^{n_m}$ be a full-dimensional compact convex set, and let $n = \sum_m n_m$. The public feasible set $\mathcal{X} = \{x \in \prod_m D_m : Ax = b\}$ is nonempty. The protocol permits any $x \in \mathcal{X}$ as its terminal allocation after a finite sequence of local queries. Selecting that allocation consumes one intervention and no additional private response. The common query limits and budgets permit the no-query policy. Both policy classes have the same terminal choices. The public model specifies the functions $f_{m,0}, h_m$, the positive amplitude α, and the uniform prior on a hidden sign σ. Party m observes its own objective $f_{m,\sigma} = f_{m,0} + \sigma\alpha h_m$. The ordinary interface returns only the optimizer in Eq. (11). It does not expose private solver traces. Assume $f_{m,0}$ and h_m are differentiable near D_m. For every allowed $q \in \mathcal{Q}_m$, assume $\psi_m(\cdot; q)$ is differentiable and convex near D_m. Both perturbed objectives are μ-strongly convex on D_m. This makes the local and shared minimizers unique. Define the baseline response region

$$\mathcal{R}_m = \{R_{m,0}(q) : q \in \mathcal{Q}_m\}, \qquad R_{m,0}(q) = \arg\min_{u \in D_m} \{f_{m,0}(u) + \psi_m(u; q)\}. \tag{187}$$

The complete parameter range in $\mathcal{Q}_m$ matters: for a proximal query it includes both the center and penalty, as well as any linear price.

Lemma L.1 (Preservation of the response map)**.** *For $r = R_{m,0}(q)$, both perturbed responses equal r if and only if Eq.* (13) *holds. In particular, $\nabla h_m = 0$ on $\mathcal{R}_m$ preserves the entire response map.*

Proof. Fix m, q and write $r = R_{m,0}(q)$. Convex optimality on D_m gives $\langle \nabla f_{m,0}(r) + \nabla\psi_m(r; q), u - r\rangle \geq 0$ for every $u \in D_m$. For each perturbed problem, this inequality is equivalent to $-\nabla f_{m,0}(r) - \nabla\psi_m(r; q) - \sigma\alpha\nabla h_m(r) \in N_{D_m}(r)$. It is necessary and sufficient by convex optimality (Bertsekas, 2012). Strong convexity makes the minimizer unique. If the perturbation gradient vanishes, the original inequality remains unchanged. At a boundary, a nonzero perturbation can also preserve membership in the normal cone. □

Couple the random tapes of a transcript policy under the two signs. At the first query the histories agree. Inductively, equal histories produce the same next query, the lemma gives equal responses, and the public transitions give equal subsequent histories. Every finite transcript, including chosen query locations and public states, has the same law. An adaptive stopping decision cannot distinguish the signs either. If objective values are also returned, the stronger condition $h_m = 0$ on $\mathcal{R}_m$ preserves those values. Zero gradient alone does not assert this extension.

L.3 Exact Value and Decision Relevance

Proof of Eq. (14). For a transcript policy, let X denote its terminal allocation, whose law is common to the two signs. Since $(F_+ + F_-)/2 = F_0$, its Bayes loss is $\frac{1}{2}\mathbb{E}[F_+(X) + F_-(X)] = \mathbb{E}[F_0(X)] \geq$

$F_0(x_0^\star)$. The constant feasible allocation $x_0^\star$ attains equality without a query. The instance-aware planner chooses $x_\sigma^\star$, proving the exact gap identity. For each sign, strong convexity and constrained optimality imply

$$F_\sigma(x) - F_\sigma(x_\sigma^\star) \geq \frac{\mu}{2}\|x - x_\sigma^\star\|^2 \qquad (x \in \mathcal{X}). \tag{188}$$

Averaging this inequality at $x = x_0^\star$ and using $\|x-a\|^2+\|x-b\|^2 = 2\|x-(a+b)/2\|^2+\|a-b\|^2/2$ gives the lower bound. □

The exact zero-gap criterion is valid at boundary optima as well. Each summand $F_\sigma(x_0^\star) - F_\sigma(x_\sigma^\star)$ is nonnegative, so their average is zero exactly when $x_0^\star$ minimizes both objectives. Differentiable convex optimality makes this equivalent to

$$-\nabla F_0(x_0^\star) - \sigma\alpha\nabla h(x_0^\star) \in N_{\mathcal{X}}(x_0^\star), \qquad \sigma = \pm 1. \tag{189}$$

No local-interiority condition is used here. The interior assumption is needed only for the following simpler projection-gradient characterization. Let $Z \in \mathbb{R}^{n\times r}$ have orthonormal columns spanning $\ker A$, where r is the dimension of the feasible direction space. If $r = 0$, the feasible set is a singleton and the gap is zero, below, take $r > 0$. Since $x_0^\star$ is interior to the local domains, $Z^\top \nabla F_0(x_0^\star) = 0$. If $g = Z^\top \nabla h(x_0^\star) = 0$, both perturbed gradients at $x_0^\star$ are normal to $Ax = b$. Convexity then makes $x_0^\star$ their common unique minimizer and $G(\alpha) = 0$. Conversely, suppose $g \neq 0$. Neither perturbed objective is stationary along $\mathcal{X}$ at $x_0^\star$. If the two objectives shared a minimizer x, their average F_0 would also be minimized at x, forcing $x = x_0^\star$. This is impossible. Their minimizers differ, so Eq. (14) implies a strictly positive gap.

L.4 The Small-Perturbation Coefficient

Assume strong convexity holds for $|t| \leq \bar{\alpha}$ in $F_t = F_0 + th$, and the functions are twice continuously differentiable near $x_0^\star$. For small t, continuity of the minimizer keeps $x_t^\star$ inside the local domains. Write $x_t^\star = x_0^\star + Zy(t)$. Stationarity is

$$Z^\top \nabla F_0(x_0^\star + Zy(t)) + tZ^\top \nabla h(x_0^\star + Zy(t)) = 0. \tag{190}$$

At $t = 0$, $y = 0$, the derivative with respect to y is the positive-definite matrix $H_T = Z^\top \nabla^2 F_0(x_0^\star) Z$. The implicit-function theorem gives $y'(0) = -H_T^{-1} g$, $x_t^\star = x_0^\star - tZH_T^{-1}g + o(t)$. Let $v(t) = F_t(x_t^\star)$. Since the tangent component of $\nabla F_t(x_t^\star)$ is zero, differentiation yields $v'(t) = h(x_t^\star)$ and $v''(0) = -g^\top H_T^{-1} g$. Therefore $v(t) = v(0) + th(x_0^\star) - \frac{t^2}{2} g^\top H_T^{-1} g + o(t^2)$. Substituting $t = \alpha$ and $t = -\alpha$ into the exact gap identity cancels the linear terms and proves Eq. (15). The coefficient is independent of the choice of orthonormal basis Z.

L.5 Active Polyhedral Constraints

Suppose $\mathcal{X}$ is a fixed compact polyhedron, F_0, h are C^2 near $\mathcal{X}$, and $F_0 + th$ has a common strong-convexity constant $\mu_0 > 0$ for sufficiently small $|t|$. Write $x_0 = x_0^\star$, $b_0 = \nabla F_0(x_0)$, $H = \nabla^2 F_0(x_0)$, and $a = \nabla h(x_0)$. Define the tangent and critical cones by

$$T_{\mathcal{X}}(x_0) = \mathrm{cl}\{t(x - x_0) : t \geq 0, x \in \mathcal{X}\}, \qquad \mathcal{K} = T_{\mathcal{X}}(x_0) \cap \{v : b_0^\top v = 0\}. \tag{191}$$

For $\sigma \in \{-1, 1\}$, let $m_\sigma = \min_{v\in\mathcal{K}} \left\{\frac{1}{2} v^\top H v + \sigma a^\top v\right\}$. Under the same response-invisibility condition, the exact gap satisfies

$$G(\alpha) = -\frac{\alpha^2}{2}(m_+ + m_-) + o(\alpha^2), \qquad \alpha \downarrow 0. \tag{192}$$

The critical-cone quadratic programs are standard tools of directional sensitivity analysis (Shapiro, 1985). Here they quantify the information loss.

Proof. Let $d_\alpha = x_{\sigma\alpha}^\star - x_0$. The two variational optimality inequalities, strong monotonicity of ∇F_0, and boundedness of ∇h imply $\mu_0\|d_\alpha\|^2 \leq \alpha\|\nabla h(x_{\sigma\alpha}^\star)\|\|d_\alpha\|$. Hence $d_\alpha = O(\alpha)$. With $v_\alpha = d_\alpha/\alpha$, Taylor expansion gives

$$F_{\sigma\alpha}(x_{\sigma\alpha}^\star) - F_0(x_0) - \sigma\alpha h(x_0) = \alpha b_0^\top v_\alpha + \alpha^2\left(\tfrac{1}{2} v_\alpha^\top H v_\alpha + \sigma a^\top v_\alpha\right) + o(\alpha^2). \tag{193}$$

The first term is nonnegative. Comparing with the feasible point x_0 shows $b_0^\top v_\alpha = O(\alpha)$. Every accumulation point of this bounded sequence therefore belongs to $\mathcal{K}$. For a fixed $v \in \mathcal{K}$, polyhedrality makes $x_0 + \alpha v$ feasible for all sufficiently small positive α. Substitution gives an upper limit of the normalized value difference at most m_σ. Along a convergent subsequence of v_α, the nonnegative first term and membership of the limit in $\mathcal{K}$ give a lower limit at least m_σ. Thus the value is $F_0(x_0) + \sigma\alpha h(x_0) + \alpha^2 m_\sigma + o(\alpha^2)$. Averaging the two signs in Eq. (14) proves the claim. □

If $\mathcal{K} = \ker A$, each minimum is $-\frac{1}{2} g^\top H_T^{-1} g$, recovering Eq. (15). At a boundary, the minimizing directions need not be opposites. Treating every active inequality as an equality would incorrectly discard one-sided feasible directions. Curved constraints are outside this expansion because their curvature can contribute additional terms.

L.6 A Continuous Saturation Example

Take two local domains $[0, 1]$, the shared constraint $x_1 + x_2 = 1$, and $f_{1,\theta}(u) = \frac{1}{2}u^2 + (1 + \theta)u$, $f_2(u) = \frac{1}{2}u^2$, $\theta \sim \mathrm{Unif}[-a, a]$, $0 < a < 1$. Ordinary price queries use $\lambda \in [0, 1]$. Both parties return zero for every such price and every θ: the lower box boundary absorbs the nonzero change in the first party's gradient. This uses ordinary quadratic objectives, with no perturbation region spliced into them. Writing $x_1 = z$, $x_2 = 1 - z$ reduces the terminal objective to $F_\theta(z) = \frac{1}{2} + z^2 + \theta z$, $z \in [0, 1]$. Its optimum is $z_\theta^\star = \max\{-\theta/2, 0\}$. The prior-mean optimum is $\overline{z} = 0$, at the boundary. Direct integration yields $G_\nu = \mathbb{E}\left[\frac{\theta^2}{4}\mathbf{1}\{\theta < 0\}\right] = \frac{a^2}{24}$. For the binary endpoints $\theta = \pm\alpha$, $\mathcal{K} = \{(t, -t) : t \geq 0\}$, $m_+ = 0$, and $m_- = -1/4$. Therefore $G(\alpha) = \alpha^2/8$, exactly matching Eq. (192). A continuous interval of distinct optimal allocations cannot be identified exactly by a deterministic fixed-length finite-bit message. Approximate optimality can instead be checked by Proposition 5.2.

L.7 Constructing Invisible Perturbations

Assume the full baseline allocation satisfies $x_0^\star \in \mathrm{int}(\prod_j D_j)$, and that some block $x_{m,0}^\star$ has positive distance from $\overline{\mathcal{R}_m}$. Choose a ball of radius r_0 about that block disjoint from $\overline{\mathcal{R}_m}$ and contained in the interior of D_m. Let χ be a smooth function supported in this ball and equal to one in a smaller ball about its center. For a vector $v \in \mathbb{R}^{n_m}$, define $h_m(u) = \chi(u)\, v^\top (u - x_{m,0}^\star)$, $h_j = 0 \quad (j \neq m)$. Both h_m and its gradient vanish near every ordinary response, while $\nabla h_m(x_{m,0}^\star) = v$. Choose v so that the corresponding block vector has a nonzero projection on $\ker A$. Such a choice exists whenever a direction in $\ker A$ changes that block. Since the full baseline allocation is interior to the local domains, Corollary 4.3 then makes the perturbation decision relevant for every admissible positive amplitude. The smooth compactly supported function has a bounded Hessian, say $\|\nabla^2 h_m\| \leq L_h$. If the baseline local objectives are μ_0-strongly convex, choosing $0 < \alpha$ with $\alpha L_h < \mu_0$ keeps both signs strongly convex with constant at least $\mu_0 - \alpha L_h$. This gives a family with identical responses and a positive gap. The baseline Hessians can be non-diagonal, and A can have several rows. At a boundary baseline, decision relevance must instead be checked using the terminal normal-cone criterion in Corollary 4.3. Appendix M uses a globally bounded-curvature ridge in place of a compactly supported perturbation, with the same response-preservation condition.

L.8 Zero-Mean Families and Posterior Decoding

Let $F_\theta = F_0 + \alpha \sum_{j=1}^p \theta_j h_j$ for a public prior with $\mathbb{E}\theta = 0$. Assume the envelope and measurability conditions of Theorem 4.2. Finite support is not required for the value identity. Each h_j is a sum of local functions whose gradients vanish on their baseline response regions. Assume every local objective is strongly convex. The same optimality argument gives identical ordinary response maps for all θ, and hence

$$V_{\mathrm{tr}} = \min_{x \in \mathcal{X}} F_0(x), \qquad G(\alpha) = \min_{x \in \mathcal{X}} F_0(x) - \mathbb{E}_\theta \min_{x \in \mathcal{X}} F_\theta(x). \tag{194}$$

Let B_T have columns $Z^\top \nabla h_j(x_0^\star)$ and let $C_\theta = \mathbb{E}[\theta\theta^\top]$. If $x_0^\star$ is interior to the local domains, the expansion under the conditions below is $G(\alpha) = \frac{\alpha^2}{2}\,\mathrm{tr}(H_T^{-1} B_T C_\theta B_T^\top) + o(\alpha^2)$. Finite support justifies averaging the remainders. A sufficient continuous-prior alternative is bounded support

$\|\theta\| \le M$, finitely many C^2 functions h_j, and a common positive strong-convexity lower bound for $|\alpha| \le \alpha_0$. Strong convexity bounds $\|x^\star_{\alpha,\theta} - x^\star_0\|$ by a constant times $|\alpha|\|\theta\|$, uniformly on that support, so all small-amplitude solutions remain in one interior neighborhood. Taylor remainders of the finitely many C^2 functions are uniform there. Solving tangent stationarity gives $y_{\alpha,\theta} = -\alpha H_T^{-1} B_T \theta + o(\alpha)$ uniformly in θ. Substitution into the objective yields a uniform $o(\alpha^2)$ remainder, which can be averaged. For unbounded parameter laws, the value identity remains available under its envelope assumption, but this expansion requires additional integrable remainder control. For any observed record W, including additional local messages, the common decision rule is

$$x_B(W) = \arg\min_{x \in \mathcal{X}} \mathbb{E}[F_\theta(x) \mid W] = \arg\min_{x \in \mathcal{X}} \left\{ F_0(x) + \alpha \sum_j \mathbb{E}[\theta_j \mid W] h_j(x) \right\}. \tag{195}$$

Thus information conditions can share a solver and differ only in their posteriors. A constant message leaves the posterior unchanged. Locally computable messages identifying the private sign coordinates recover the instance-aware value when all coordinates are known. Their cost grows with the number of coordinates transmitted. The two-instance one-bit construction is a special case. For an allocation rule $\widehat{x}(W)$, conditioning on W gives the expected-loss decomposition

$$\mathbb{E}[F_\theta(\widehat{x}(W)) - F_\theta(x^\star_\theta)] = \mathbb{E}[F_\theta(x_B(W)) - F_\theta(x^\star_\theta)] + \mathbb{E}[F_\theta(\widehat{x}(W)) - F_\theta(x_B(W))]. \tag{196}$$

The second term is nonnegative because $x_B(W)$ minimizes the conditional expected objective. This is an expectation statement: that term need not be nonnegative on each private instance.

L.9 A Clearing Price Rules Out This Obstruction

For a local price response $r_m = \arg\min_{x_m \in D_m}\{f_m(x_m) + \lambda^\top A_m x_m\}$, suppose $\sum_m A_m r_m = b$. For any $x \in \mathcal{X}$, summing local optimality inequalities yields $\sum_m f_m(r_m) + \lambda^\top b \le \sum_m f_m(x_m) + \lambda^\top b$. Hence r is globally optimal. If two strongly convex instances have identical response maps on an allowed price domain containing such a price, they share the same unique optimal allocation. This is the usual dual-optimality implication (Boyd & Vandenberghe, 2004). It characterizes sufficient interface coverage. It does not assert that a finite number of queries locates an unknown clearing price exactly.

L.10 A Noisy Extension

For completeness, consider a different observation model with Gaussian noise: a query q yields $Y = R_\sigma(q) + \xi$, where $\xi \sim \mathcal{N}(0, \sigma_\xi^2 I)$ independently at each contact. Suppose the two means differ by at most δ_R at every admissible query and history. A common adaptive policy using at most Q contacts has $\mathrm{KL}(P_+^{\pi,W} \| P_-^{\pi,W}) \le \frac{Q\delta_R^2}{2\sigma_\xi^2}$, $\mathrm{TV}(P_+^{\pi,W}, P_-^{\pi,W}) \le \min\{1, \delta_R\sqrt{Q}/(2\sigma_\xi)\}$. The first inequality follows from the conditional KL chain rule: query and public-transition kernels are shared, and each Gaussian response contributes at most $\delta_R^2/(2\sigma_\xi^2)$. Pinsker's inequality gives the second. These are standard interactive testing tools (Chen et al., 2024). If both terminal objectives are μ-strongly convex and their optima have distance Δ_x, integrate their excess losses against the common part of the two record laws, of mass $1 - \mathrm{TV}$. Conditional on the same record, the policy uses the same terminal action kernel. The squared-distance identity from Appendix L.3 therefore gives $\mathrm{Gap}_{\mathrm{rec}} \ge \frac{\mu\Delta_x^2}{8}\left(1 - \frac{\delta_R\sqrt{Q}}{2\sigma_\xi}\right)_+$. Noise is part of this statement's observation model. Unequal deterministic replies at unlimited precision need not have small total variation, regardless of their Euclidean distance. The experiments in Appendix M concern exact aliasing and do not test this noisy extension.

M Controlled Experiments on Query-Invisible Perturbations

This appendix specifies the direct tests in Section 6.1. The objective families satisfy Theorem 4.2. The experiments measure its exact value identity, sensitivity prediction, and distinction between invisible and decision-relevant changes. All information conditions share a posterior-optimal terminal decoder. The task is separate from the hourly-separable power task in Appendix G.

M.1 Public Profiles and Coupled Baseline

We use actual-load columns from the OPSD Time Series package, version 2020-10-06 (Open Power System Data, 2020), for Portugal, Slovenia, Finland, Switzerland, Slovakia, Norway, Sweden, and Latvia, in that order. Each party has 24 hourly observations in UTC. The four complete dates are 2017-07-01, 2017-07-04, 2017-07-15, and 2017-07-19, the first four entries in the existing development-date list. They are used only as public task profiles. We do not inspect the reserved evaluation outcomes or select dates by the resulting gaps. The raw file is `time_series_60min_singleindex.csv`. Its version, provider, and source citation are shared with Appendix G.1. For party m and hour h, normalize its load L_{mh} by its mean over that day:

$$r_{mh} = \mathrm{clip}_{[0.5,1.5]}\left(\frac{L_{mh}}{24^{-1}\sum_{j=1}^{24} L_{mj}}\right), \qquad p_{mh} = 0.6 + 0.3 r_{mh}. \tag{197}$$

The local domain is $D_m = \prod_h [0, p_{mh} + 2]$. With $M = 8$, public demand is

$$b_h = \sum_m p_{mh} + M\left(0.35 + \frac{0.1}{M}\sum_m r_{mh}\right), \qquad \mathcal{X} = \{x \in \prod_m D_m : \textstyle\sum_m x_{mh} = b_h \ \forall h\}. \tag{198}$$

We generate $d_{mh} \sim \mathrm{Unif}[1, 1.8]$ and set

$$H_m = \mathrm{diag}(d_{m1}, \ldots, d_{m,24}) + \gamma L^\top L, \qquad f_{m,0}(x_m) = \frac{1}{2}(x_m - p_m)^\top H_m (x_m - p_m), \tag{199}$$

where $(Lv)_h = v_{h+1} - v_h$. Thus the temporal term penalizes changes in deviations from the preferred profile. Adjacent hours are coupled within each party, and every hour is coupled across parties by Eq. (198). The baseline Hessian lower bound is $\mu_0 = \min_m \lambda_{\min}(H_m) \geq 1$. The equality-constrained baseline solution is

$$\eta_0 = \left(\sum_m H_m^{-1}\right)^{-1}\left(b - \sum_m p_m\right), \qquad x^\star_{m,0} = p_m + H_m^{-1}\eta_0. \tag{200}$$

All returned baseline and perturbed solutions satisfy the box constraints strictly. Across the 840 signed-pair evaluations below, the smallest terminal distance to a box face is 1.018. Consequently, these equality solutions are also the box-constrained optima.

M.2 The Invisible Ridge and the Full Query Domain

Let w be a public unit vector obtained by normalizing the positive vector with coordinates $1 + 0.8\sin\left(\frac{2\pi(h-1)}{23} + \omega\right)$, $\omega \sim \mathrm{Unif}[0, 2\pi]$. For width $a = 0.5$, use the twice continuously differentiable ridge

$$\phi(s) = \begin{cases} 0, & s \leq 0, \\ s^3/(6a), & 0 < s < a, \\ s^2/2 - as/2 + a^2/6, & s \geq a. \end{cases} \tag{201}$$

Its second derivative lies in $[0, 1]$. Put $\ell = 0.02$ and define the public thresholds

$$t_m = w^\top p_m + \frac{\sqrt{24}\ell}{\mu_0} + 0.05, \qquad h_m(x_m) = c_m \phi(w^\top x_m - t_m). \tag{202}$$

The ordinary query domain consists of every party-specific vector $\lambda \in [-\ell, \ell]^{24}$, with response $R_{m,\sigma}(\lambda) = \arg\min_{u \in D_m}\{f_{m,0}(u) + \sigma\alpha h_m(u) + \lambda^\top u\}$. There is no free proximal center or additional hidden query parameter. For every allowed price, the baseline response is $r_m = p_m - H_m^{-1}\lambda$. It lies inside D_m, since $\|H_m^{-1}\lambda\|_2 \leq \sqrt{24}\ell/\mu_0 < 0.098$ and $p_{mh} \geq 0.75$. Also, $w^\top r_m - t_m \leq \|H_m^{-1}\lambda\|_2 - \sqrt{24}\ell/\mu_0 - 0.05 \leq -0.05$. Thus $h_m(r_m) = 0$ and $\nabla h_m(r_m) = 0$ over the whole price cube. When $|c_m| \leq 1$ and $\alpha < \mu_0$, each perturbed objective is at least $(\mu_0 - \alpha)$-strongly convex, so $R_{m,\sigma}(\lambda) = r_m$ under both signs. This proves full-domain response equality, not just agreement on the finite queried sample. The implementation returns this common analytic optimizer using the same float64 arithmetic under both signs. It sends no private solver iterates or stopping diagnostics. To test the assembled transcript, an adaptive policy selects its next party and price from the preceding response and a shared random tape. At depths $Q \in \{0, 8, 32, 128\}$, every recorded party, price, and response agrees bit-for-bit between signs. The maximum observed difference is zero.

M.3 Relevant and Normal Perturbations

For the decision-relevant family, independently draw $u_m \sim \mathrm{Unif}[0.5, 1]$ and set $c_m = (-1)^{m-1} u_m / \max_j u_j$. The private objectives are $f_{m,\sigma} = f_{m,0} + \sigma \alpha c_m \phi(w^\top x_m - t_m)$ with equiprobable $\sigma \in \{-1, 1\}$. All baseline matrices, ridge functions, coefficient magnitudes, and the prior are public. The sign is hidden. For the normal control, let $s_m^0 = w^\top x_{m,0}^\star - t_m$ and set

$$c_m^{\mathrm{normal}} = \frac{\min_j \phi'(s_j^0)}{\phi'(s_m^0)}. \tag{203}$$

All denominators are positive in the declared tasks. At the baseline optimum, every party's perturbation gradient is the same vector, $\min_j \phi'(s_j^0) w$. This vector is normal to $\sum_m x_m = b$, so $Z^\top \nabla h(x_0^\star) = 0$ and $x_0^\star$ minimizes both signed objectives. The normal and relevant constructions share the public profiles, feasible set, baseline Hessians, ridge, and query domain. We use seeds 11, 23, 37, 53, and 71 with NumPy's `default_rng`, drawing d_{mh}, then ω, then u_m. The three couplings are $\gamma = 0.25, 1, 4$. The seven amplitudes are $\alpha = 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, 0.6$. The four days, five seeds, and three couplings give 60 public tasks. Seven amplitudes and two perturbation types give 840 signed-pair evaluations. No coefficient is fitted to the measured gaps.

M.4 Common Decoder and Additional Information

For a public finite family and prior, all conditions choose $x_B(W)$ from Eq. (195). Ordinary transcripts leave the prior unchanged, so the decoder returns $x_0^\star$. The reported loss is the raw objective difference, averaged with the public prior. No instance-dependent normalization is applied. For the binary family, three convex solves compute the exact Bayes gap in Eq. (14).

A locally computed bit. At the public point $x_{m,0}^\star$, party m evaluates

$$d_m^{\mathrm{loc}} = w^\top \left[\nabla f_{m,\sigma}(x_{m,0}^\star) - \nabla f_{m,0}(x_{m,0}^\star) \right] = \sigma \alpha c_m \phi'(s_m^0). \tag{204}$$

It sends $C_m = \mathbf{1}\{c_m d_m^{\mathrm{loc}} > 0\}$, which equals $\mathbf{1}\{\sigma = 1\}$. In the binary experiment party 1 sends this bit. The constant-message control performs the same directional-derivative calculation but sends zero regardless of its value. Both add one transmitted bit and one local derivative evaluation. Only one changes the posterior.

An expanded price query. We additionally allow a query to party 1 at the public price $\lambda^\dagger = -H_1(x_{1,0}^\star - p_1)$. This price is outside the ordinary cube and moves the baseline local optimum to $x_{1,0}^\star$, where the ridge has nonzero derivative. The two private replies differ. The common decoder compares the reply to the two public-family responses and updates its posterior. This condition changes query permission and adds a 24-dimensional optimizer message. It is not an algorithm within the original price domain.

Sixteen coupled instances. We replace the single sign by four independent equiprobable signs θ_j, sharing sign j between parties $2j - 1$ and $2j$. The coefficient at party m is then $\alpha c_m \theta_{\lceil m/2 \rceil}$. All 16 instances use the same 24 balance constraints and temporal Hessians. Parties 1, 3, 5, 7 can each compute their own sign by Eq. (204). After $b_C = 0, \ldots, 4$ bits, in that public order, the decoder minimizes the posterior-mean objective over the signs still possible. Four constant bits provide the matched-message control. This family is evaluated at $\alpha = 0.3$ on all 60 public tasks. The table counts responses and contacts. A price-query request also contains a party identifier and 24 float64 coordinates. A certificate request contains its party identifier and the public statistic identifier. These request formats are identical within each matched-message comparison. The instance-aware value is an information benchmark, not a claim of zero-cost transmission of the private objectives.

M.5 Solvers and Numerical Results

The shared decoder uses equality-constrained Newton steps. Each local Hessian is $H_m + \alpha c_m \phi''(s_m) w w^\top$ (with the appropriate posterior coefficient), so a rank-one inverse update and a 24-dimensional Schur-complement solve compute the step. An Armijo line search keeps iterates inside the public boxes. We stop when the maximum absolute component of the gradient projected

Table 28: Additional resources in the information comparisons. All conditions use the same terminal solver. Ordinary and expanded queries return float64 optimizer coordinates. Certificates are literal binary symbols.

Information	Extra contacts	Response bits	Local operation per contact
Q ordinary queries	Q	$1536Q$	One 24-dimensional quadratic solve using the public matrix factorization
One constant bit	1	1	One directional-derivative evaluation, value discarded
One local sign bit	1	1	The same directional-derivative evaluation
Expanded-price query	1	1536	One strongly convex local solve at the additional price
Four local sign bits	4	4	One derivative evaluation at each of four parties
Four constant bits	4	4	The same four evaluations, values discarded

onto $\sum_m x_m = 0$ is at most 2×10^{-10}, with an 80-step limit. Every returned allocation is checked for balance error at most 2×10^{-9} and distance greater than 10^{-5} from all box faces. Across the 840 signed-pair evaluations, the maximum balance error is 5.33×10^{-15} and the maximum projected-gradient residual is 3.56×10^{-15}. Their terminal solutions remain in the quadratic part of the ridge, so at most one Newton update is needed from $x_0^\star$. The Hessians and shared Schur complement remain non-diagonal. The small-amplitude coefficient is computed directly from the baseline constrained Hessian, without using the perturbed solutions. All 420 relevant pairs satisfy the strong-convexity lower bound in (14). The smallest difference between the exact gap and that bound is 2.99×10^{-7}. Every normal pair has identical optimal allocations, with maximum absolute computed gap 3.56×10^{-15}. In the binary information comparison, the minimum expanded-query response separation is 0.546. Local bits and expanded queries yield zero remaining gap to the reported solver precision. The ordinary and constant-bit conditions both retain mean gap 0.438434748 at every evaluated ordinary depth.

Table 29: Relevant perturbations over all 60 public tasks per amplitude. Error is the maximum relative difference between the exact gap and Eq. (15)'s leading term. Ranges span the fixed task grid, not sampling confidence intervals.

α	Mean gap	Gap range	Maximum relative error
0.001	$4.86067\,10^{-6}$	$[4.00626, 5.40358]\,10^{-6}$	$4.90\,10^{-8}$
0.003	$4.37460\,10^{-5}$	$[3.60563, 4.86322]\,10^{-5}$	$4.36\,10^{-7}$
0.010	$4.86068\,10^{-4}$	$[4.00627, 5.40359]\,10^{-4}$	$4.85\,10^{-6}$
0.030	0.00437470	[0.00360577, 0.00486333]	$4.36\,10^{-5}$
0.100	0.0486184	[0.0400804, 0.0540499]	$4.86\,10^{-4}$
0.300	0.438435	[0.362048, 0.487495]	0.004515
0.600	1.76722	[1.46867, 1.96572]	0.020294

Table 30: Mean Bayes gaps in the 16-instance family at $\alpha = 0.3$. Local bits identify successive private sign coordinates. Constant bits leave the prior unchanged.

Certificate bits	0	1	2	3	4
Local sign bits	0.438642	0.318287	0.205159	0.087085	0
Constant bits	0.438642	0.438642	0.438642	0.438642	0.438642

Numerical checks. Eight tests check the ridge derivatives, the common equality solver against an independent SLSQP implementation on a smaller problem, every price-cube corner of that smaller problem, adaptive transcripts, the normal control, the sensitivity coefficient, local message recovery, and the marginal-response multiplier convention in Appendix K. The experiments and tests were run on the remote server with NumPy 2.2.6 and SciPy 1.15.3. These are controlled finite-family

information experiments. The Gaussian-noise extension in Appendix L.10 is a theoretical result rather than an additional empirical condition.

N Record Compatibility in the Original Resource Task

This experiment uses the resource-allocation objectives of Appendix G, without the ridge perturbation of Appendix M. It asks what uncertainty remains after a prescribed response record, and separately evaluates the list-free certificate of Proposition 5.2. The dates, parameter envelope, message formats, search, and reporting threshold were specified before generating the results. No predictor is trained.

N.1 Tasks, Information, and Response Records

We use 29 July, 4 August, 6 August, and 7 August 2017, the fifth through eighth dates of the existing development partition for the response-protocol study. The five objective seeds are 1709, 1721, 1723, 1733, and 1741. The conflict strengths are 0.75 and 1.0. Their Cartesian product gives 40 tasks. Normalization uses the original 2015–2016 training interval. The eight parties correspond to Portugal, Slovenia, Finland, Switzerland, Slovakia, Norway, Sweden, and Latvia. Their resource counts are $(5, 5, 8, 8, 3, 5, 5, 4)$, totaling 43 resources over 24 hours. Data provenance and the original parameter generator are given in Appendix G. For resource r of party m at hour h, the cost remains $f_{mrh}(x) = \frac{1}{2}c_{mrh}(x - p_{mrh})^2 + \omega_{mrh}|x - p_{mrh}|$, $0 \le x \le \bar{x}_{mrh}$. Let $V_m(y_m)$ be the minimum sum of these costs subject to the party's hourly resource totals being y_m. Write $\Phi(y) = \sum_m V_m(y_m)$, $U_{mh} = \sum_r \bar{x}_{mrh}$, and $\mathcal{Y} = \{y : 0 \le y \le U,\ \sum_m y_{mh} = d_h\}$. All reported losses in this appendix are raw excess values of Φ, not the normalized regret used in the earlier intervention experiments.

Fixed ordinary protocol. At query center z_m and scaled dual u_m, party m returns the resource total minimizing $\sum_{r,h} f_{mrh}(x_{rh}) + \frac{1}{2}\left\|\sum_r x_r - z_m + u_m\right\|^2$. The penalty is one, and the local multiplier solve uses the original 50 bisection steps, brackets, equality branch, and midpoint return. The ordinary message contains only 24 resource totals. The initial public allocation is the box-and-balance projection of the preferred totals, with zero duals. All eight parties answer once at that initial state. Their responses are projected and the duals are updated. Subsequent contacts cycle through party indices $0, \dots, 7$, updating the queried cached response, projecting the cached responses plus duals, and updating the duals after each contact. We retain prefixes with $Q = 0, 8, 32, 128$ additional contacts. Thus $Q = 0$ already contains eight responses. We evaluate two predeclared message formats. Float64 records require equality of all 64-bit floating-point representations. Fixed-point records use signed 16-bit integers with grid width $\Delta = 2^{-12}$, nearest-even rounding, and overflow rejection. The coordinator uses the decoded grid values in its updates. There is no post hoc distance tolerance for matching records. The two formats can induce different subsequent query locations.

N.2 Record-Conditioned Decision Loss

The original generator satisfies $.04a_r \le \omega_{mrh} \le .32a_r$, $a_r = .65 + .3(r + 1)$, $r = 0, \dots, R_m - 1$. We fix preferences, capacities, curvatures, and demand, and allow the penalties to vary independently within this envelope. These fixed quantities are granted as additional public information for this diagnostic. The resulting class $\mathfrak{F}$ is a parameter envelope: it is not the joint distribution induced by the original signal-to-cost generator. A compatible alternative supports a worst-case uncertainty statement within this class. For a record W containing query locations, public states, and encoded messages, define $\mathfrak{F}(W) = \{\Phi \in \mathfrak{F} : \Phi \text{ reproduces every encoded response in } W\}$, $\mathcal{R}(W) = \inf_{y \in \mathcal{Y}} \sup_{\Phi \in \mathfrak{F}(W)} [\Phi(y) - \min_{\mathcal{Y}} \Phi]$. The deterministic public update then reproduces the same trajectory for any member of $\mathfrak{F}(W)$. This inverse-optimization perspective uses observed optimal behavior to constrain objectives (Bertsimas et al., 2015). The quantity of interest here is the difference between compatible objectives' decisions. The aggregate objectives share the strong-convexity bound $\mu_{\text{agg}} = \min_{m,h} \left(\sum_r c_{mrh}^{-1}\right)^{-1}$. Indeed, for any resource perturbation, weighted Cauchy–Schwarz bounds its quadratic cost below by this harmonic curvature times the squared change in its total. Partial minimization and summation preserve the resulting strong-convexity inequality. For any verified compatible pair with optima $y_0^\star, y_1^\star$,

$$\mathcal{R}(W) \ge \frac{\mu_{\text{agg}}}{8}\|y_0^\star - y_1^\star\|^2. \tag{205}$$

Table 31: Original-task record diagnostic. Each row contains the same 40 tasks. Above and below use the prespecified raw-loss threshold 10^{-4}. Unresolved records satisfy neither test. Mean lower bounds include zero when the search finds no separated pair.

Message	Extra Q	Above	Below	Unresolved	Mean lower bound
Float64	0	22	0	18	$1.4371\,10^{-2}$
Float64	8	0	0	40	0
Float64	32	0	0	40	0
Float64	128	0	0	40	0
Fixed-point	0	40	0	0	$9.0552\,10^{-2}$
Fixed-point	8	40	0	0	$1.0981\,10^{-2}$
Fixed-point	32	0	0	40	$4.4619\,10^{-6}$
Fixed-point	128	0	0	40	$5.2567\,10^{-8}$

The proof is the two-point quadratic-growth argument of Eq. (186), applied after fixing W.

Candidate search. For each party, we identify resource coordinates that equal their preference in every recorded local solve. Such a coordinate can retain a range of penalties without changing the exact proximal optimum. Its candidate lower endpoint is the larger of the public penalty lower bound and the largest recorded absolute multiplier bound plus 10^{-9}. Its upper endpoint is the public upper bound. Coordinates with an empty interval, or not fixed at their preference, retain their original penalty. We form low and high variants for each party. Each variant is retained only if rerunning the actual 50-step response implementation reproduces every message for that party, exactly in the declared encoding. We test single-party variants, all-low and all-high combinations, and two alternating combinations, together with the original objective: at most 21 candidates per record. Partywise matches compose because the public updates are deterministic. Private resource allocations are used only by the offline search to propose candidates. They are not additional online observations. The candidate collection is not an exhaustive description of $\mathfrak{F}(W)$. Central optima use 80 multiplier bisections. In Eq. (205), we reduce the computed separation by the sum of the two solver-error bounds, clipping at zero. For each hour the multiplier half-bracket, with a $64\epsilon_{\mathrm{mach}}(1+\max\{|\mu^-|,|\mu^+|\})$ floating-point padding, is multiplied by $\sum_r c_{mrh}^{-1}$ for each party. The Euclidean norm bounds allocation error in exact-arithmetic bisection with that padding. The results use float64 computation, not directed-rounding interval arithmetic. Exact equality of the transmitted message bits is checked separately.

Upper bounds and unresolved records. Let Φ^{lo} and Φ^{hi} set every penalty to its public lower and upper bound. Pointwise, $\Phi^{\mathrm{lo}} \le \Phi \le \Phi^{\mathrm{hi}}$ for every member of the class. Choosing a minimizer of Φ^{hi} gives $\mathcal{R}(W) \le \min_{\mathcal{Y}} \Phi^{\mathrm{hi}} - \min_{\mathcal{Y}} \Phi^{\mathrm{lo}}$. This bound applies to the entire envelope, so it remains valid for any record-compatible subset, but need not decrease with query depth. We classify a record as above threshold when its pair lower bound exceeds 10^{-4}, below threshold when the class upper bound is at most 10^{-4}, and unresolved otherwise. A failed candidate search is therefore not counted as absence of ambiguity. None of the upper bounds was small enough to establish a below-threshold case.

N.3 Results across Depth and Precision

Table 31 reports all 320 conditions. At initialization, the float64 witness rate is 55%. Its 95% percentile interval from 2,000 whole-date bootstrap resamples is $[35\%, 80\%]$. Fixed-point records have a witness above threshold in every task at depths zero and eight, yielding a degenerate empirical interval at one. These intervals describe four selected development dates, not the frequency of ambiguity in a population of electricity systems. Later zero detection rates likewise describe this incomplete search. The class-wide upper bound averages 4.6282 across the 40 tasks, much larger than the reporting threshold at every depth. Under float64 responses, the mean fraction of queried resource coordinates exactly at their preference falls from 96.06% at initialization to 73.39% at depth 128. The fraction at a box boundary rises from 0.107% to 10.30%. The fixed-point fractions are nearly identical. The search specifically exploits the preference kink in the original absolute-deviation cost.

Table 32: List-free certificate sizes. Each entry averages 40 tasks. Two certificate-message precisions are evaluated at each ordinary-record allocation. All 640 numerical bounds cover the evaluated loss, but none is at most 10^{-3} or 10^{-4}.

Ordinary message	Extra Q	Actual excess	Float64 bound	Fixed-point bound
Float64	0	0.600718	30.6506	30.6982
Float64	8	0.412677	30.0364	30.0835
Float64	32	0.034384	13.5381	13.5691
Float64	128	0.000632	12.1919	12.2217
Fixed-point	0	0.600741	30.6498	30.6973
Fixed-point	8	0.412603	29.9086	29.9557
Fixed-point	32	0.034411	13.5459	13.5772
Fixed-point	128	0.001084	16.7295	16.7651

It does not establish that box saturation caused the witnesses, nor that their equality extends beyond the recorded queries.

N.4 List-Free Certificate Evaluation

At each retained projected allocation, the coordinator requests an additional marginal-value message from every party. Each party performs the 50-step local equality solve described in Appendix F.6, returns 24 gradient coordinates, and supplies one Euclidean error bound. The coordinator combines the eight error bounds by their root sum of squares. Its computation uses only the allocation, public capacities and balance constraint, received messages, and the public curvature lower bound $\mu_{\mathrm{pub}} = \min_m \frac{.35}{\sum_{r=0}^{R_m-1} 1/(.65+.3(r+1))}$. It does not receive the private penalty array, a list of candidate objectives, or the true loss. We separately encode the gradient message in float64 and on the same fixed-point grid. The latter adds $\Delta/2$ to each coordinate's solver-error bound. The residual is computed by a 70-step scalar normal-projection solve per hour. Using the returned valid normal vector in Proposition 5.2 preserves an upper bound even if that projection is not exact. Centralized objective evaluation is used only afterward to check the reported certificate. Each certificate costs eight additional contacts and $50 \sum_m R_m \upsilon_m = 4{,}800$ local-work units, where the public per-resource multipliers are $\upsilon = (4, 4, 1, 2, 1, 2, 3, 1)$. The response payload is 12,288 gradient bits in float64 or 3,072 bits in fixed point, plus 512 bits for the eight error bounds. Sending the 192 requested allocation coordinates adds 12,288 float64 bits in either case. Message headers and transport overhead are not included. The normal-projection solve is coordinator computation, not additional private response work. These certificate queries do not alter the preceding ordinary record. Table 32 distinguishes certificate validity from usefulness as a stopping rule. The public curvature bound is conservative. Moreover, at a kink of the aggregate value function, the multiplier convention returns one particular valid subgradient. It need not be the subgradient closest to constrained stationarity. Thus a small actual loss can coexist with a large reported bound. All bounds remain above the two declared certification tolerances. The maximum balance error of the evaluated allocations is $2.67\,10^{-15}$.

N.5 Numerical Checks and Interpretation

Numerical tests check the continuous saturation example, its critical-cone coefficient, normal-cone projection, the bounded-error certificate, bitwise message comparisons, and aggregate curvature. Additional replay checks reconstruct the verified candidate pairs without changing the candidate search. The controlled study in Appendix M proves equality over a full permitted query domain and computes an exact Bayes gap under a public prior. The present diagnostic establishes a different fact: some finite records in the original task admit decision-relevant alternatives in a specified cost envelope. It neither optimizes over all query policies nor assigns posterior probability to those alternatives. This separation makes it possible to report positive witnesses and unresolved cases without interpreting a search failure as recoverability.