THE STABLE DISCRIMINANT AND ADAPTIVE GUARDS FOR AUDITING STABLE MATCHINGS

Anonymous authors
Paper under double-blind review

ABSTRACT

Auditing a proposed matching need not recover every preference sign. We derive exact Gaussian information games for scalar queries and for bilateral feedback from deployed matchings. Stability requires ruling out every candidate blocker; instability requires certifying only one. Shared comparisons explain the resulting sampling allocations, while matching feasibility changes which blocker is easiest to certify. Adaptive guards exploit the stronger rejecting endpoint and achieve finite expected deployment cost within logarithmic factors of the information bound. A separate tracker attains the asymptotic leading constant using a witness-local stopping boundary. Finally, a single design certificate combines an assignment-based optimization bound with simultaneous confidence intervals, covering both numerical accuracy and statistical uncertainty. Synthetic experiments distinguish these guarantees, and explicit markets show why auditing, finding an optimal stable matching, and learning the stable-pair graph have different information costs.

Keywords: Stable matching; adaptive auditing; fixed-confidence learning; information-theoretic complexity; matching feedback; sequential testing.

1 INTRODUCTION

Stable matching supports market design when both sides have preferences over their partners (Gale & Shapley, 1962; Gusfield & Irving, 1989; Roth, 1984; Abdulkadiroglu & Sonmez, 2003; Hatfield & Milgrom, 2005). When utilities are observed through noise, a basic deployment question is whether a supplied matching is stable. Answering this question can require much less information than learning complete preference orders or identifying every stable outcome.

The reason is local. A candidate pair blocks a matching only when *both* agents prefer each other to their assigned partners. For $\varepsilon > 0$, if their gains are $(-1, -\varepsilon)$, the first agent's rejection already rules out the pair, even as ε tends to zero. For gains $(-\varepsilon, 1)$, the small negative gain is the only protection and must be resolved. An auditor should distinguish these cases before paying to learn every sign.

We develop this distinction in three steps. First, exact Gaussian projections give the information cost of creating or destroying a blocker. The stable answer induces a universal max–min design; an unstable answer selects one witness. Under deployed-matching feedback, feasible exposures are doubly stochastic, and a half–half partner-swap experiment is optimal for a fixed witness. Second, we give two complementary algorithms: finite-confidence guards avoid redundant near ties, while information tracking attains the asymptotic optimal constant. A local stopping boundary tests only the observations relevant to a witness. Third, confidence envelopes and an assignment upper bound certify the information rate of an adaptively chosen design. The resulting guarantee separates numerical optimization error from uncertainty about the market.

Relation to prior work. Matching bandits study repeated stable outcomes and regret (Liu et al., 2020; 2021; Cen & Shah, 2021; Kong et al., 2024). Fixed-confidence work identifies a stable or optimal stable matching (Hosseini et al., 2024; Pagare et al., 2025; Athanasopoulos et al., 2026); our output certifies a prescribed deployment, with one stable answer or several possible blocking witnesses. Preference elicitation studies deterministic comparisons and partial information (Drummond & Boutilier, 2013; Rastegari et al., 2013; Aziz et al., 2020), whereas noisy ranking recovers full orders (Feige et al., 1994; Braverman & Mossel, 2008; Jamieson & Nowak, 2011; Heckel et al., 2019). We

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.

use established change-of-measure, multiple-answer tracking, and combinatorial information-game tools (Kaufmann et al., 2016; Garivier & Kaufmann, 2016; Degenne & Koolen, 2019; Jourdan et al., 2021; Kaufmann & Koolen, 2021; Poiani et al., 2026). The Gaussian star allocation is an inherited best-arm subproblem. Our contribution is the bilateral audit geometry, its feasible designs, and the guard and certification guarantees. Appendix K gives the broader comparison.

2 PRELIMINARY

This section collects the shared notation; proof-local variables are introduced in their proofs. Write $[n] = \{1, \dots, n\}$, $\Delta(S) = \{q \in \mathbb{R}_+^S : \sum_{s \in S} q_s = 1\}$, and $\Delta_k = \Delta([k])$. The symbol $\mathbf{1}$ denotes an all-ones vector (dimension determined by context), and $\mathbf{1}\{E\}$ an event indicator. Logs are natural; $\|\cdot\|_2$ and $\|\cdot\|_\infty$ are Euclidean and maximum norms. Matrix inner products sum entrywise products; $\mathrm{supp}(q) = \{p : q_p > 0\}$. Degree-one homogeneity means $F(cz) = cF(z)$ for $c > 0$.

2.1 MARKET, TARGETS, AND OBSERVATIONS

Let A and B be disjoint copies of $[n]$, $n \geq 2$. A matching is a bijection $M : A \to B$, also written $(M(1), \dots, M(n))$. The unknown utilities are $\theta = (u, v) \in \mathbb{R}^d$, $d = 2n^2$, where u_{ij} and v_{ji} are the two agents' utilities on edge (i, j). Higher utility means greater preference, denoted $\succ$; a strict profile has no within-agent ties. Its minimum preference gap is the smallest absolute difference between distinct entries in any agent's list. For a supplied matching M and $p = (i, j)$, set

$$\mathcal{P}_M = \{(i, j) : j \neq M(i)\}, \qquad P = |\mathcal{P}_M| = n(n-1),$$
$$g_p^A = u_{ij} - u_{iM(i)}, \qquad g_p^B = v_{ji} - v_{jM^{-1}(j)}.$$

Pair p blocks M if both gains are positive; $\texttt{block}(M)$ is the set of such pairs. The matching is stable if this set is empty. Write $[z]_+ = \max\{z, 0\}$, $\Delta_p^s = [-g_p^s]_+$ for $s \in \{A, B\}$, $d_p = \max_s \Delta_p^s$, and $d_* = \min_p d_p$ on a stable instance. The *strict audit domain* requires $g_p^s \neq 0$ for every p, s; ties between off-match alternatives are allowed. Unless stated otherwise, audit guarantees use this domain; structural targets require fully strict profiles.

Definition 2.1 (Targets and discriminant)**.** A stable-market target Ψ is a finite-valued function of the strict preference profile. Target Ψ_1 refines Ψ_2, written $\Psi_1 \succeq \Psi_2$, if $\Psi_2 = f \circ \Psi_1$ for some deterministic f. Its stable discriminant $\mathfrak{D}_\Psi$ consists of points whose every neighbourhood contains strict parameters with different target values; its margin is $\kappa_\Psi(\theta) = \mathrm{dist}_\infty(\theta, \mathfrak{D}_\Psi)$. The union of within-agent tie hyperplanes is $\mathfrak{H}$. These hyperplanes form the tie arrangement: its open chambers fix all preference orders, and its faces allow specified ties. Distance is measured in the maximum norm.

Here Ψ_{prof} is the full collection of preference orders, $\mathcal{S}(\theta)$ the set of all stable matchings, and $M^A(\theta)$ the stable matching every A-agent weakly prefers to every other stable matching. It is returned by A-proposing deferred acceptance: unmatched A-agents propose in preference order, and each B-agent keeps its best offer so far. The stable-pair graph is $G(\theta) = \bigcup_{L \in \mathcal{S}(\theta)} \{(i, L(i)) : i \in A\}$; the binary audit target is $\Psi_M^{\mathrm{aud}} = \mathbf{1}\{M \in \mathcal{S}\}$.

Two observation models. A scalar query observes one independent $\mathcal{N}(\theta_e, 1)$ utility, where e indexes a side-specific coordinate. Its allocation $w \in \Delta_d$ has entries w_{ij}^A, w_{ji}^B. A matching round deploys one perfect matching and observes both independent utilities on every assigned edge, returning $2n$ scalar observations. Its exposure matrix $x \in \mathcal{B}_n$ gives edge-observation probabilities, where

$$\mathcal{B}_n = \{x \in \mathbb{R}_+^{n \times n} : x\mathbf{1} = \mathbf{1},\ x^\top \mathbf{1} = \mathbf{1}\}.$$

This doubly stochastic (Birkhoff) polytope consists of mixtures of permutation matrices. A Birkhoff decomposition expresses x as such a mixture; sampling its matching with the corresponding mixture probability realizes x. Uniform exposure has $x_{ij} = 1/n$. Cyclic initialization deploys $i \mapsto 1 + ((i + k - 1) \bmod n)$ for $k = 0, \dots, n-1$.

For $p = (i, j)$, its partner edges are $e_A = (i, M(i))$ and $e_B = (M^{-1}(j), j)$, and $x_p = x_{ij}$. The set Q_p indexes $u_{ij}, u_{iM(i)}, v_{ji}, v_{jM^{-1}(j)}$. The physical-edge count $N_{ij}(t)$ is shared by two utility streams; $N_e(t)$ and $\widehat{\theta}_e(t)$ are utility counts and means, and $N(t), \widehat{\theta}(t)$ their vectors. The prefix mean

$\widehat{\theta}_{e,k}$ uses the first k observations of stream e. Hats on gains, deficits, or rates mean substitution of empirical utilities. Zero-count means are omitted until initialization. $\mathbb{P}_\theta, \mathbb{E}_\theta$ denote probability and expectation under θ.

2.2 Information, stopping, and confidence notation

For a finite answer set $\mathcal{I}$, let $\mathcal{I}^*(\theta)$ be the correct answers and $\mathrm{Alt}_a = \{\lambda : a \notin \mathcal{I}^*(\lambda)\}$, with λ in the stated parameter domain. For $0 < \delta < 1/2$, an auditor is δ-correct if it stops almost surely and returns a member of $\mathcal{I}^*(\theta)$ with probability at least $1 - \delta$ on every instance in its domain. Auditing has $\mathcal{I}^* = \{\texttt{stable}\}$ when M is stable and $\mathcal{I}^* = \texttt{block}(M)$ otherwise. With cl denoting closure, define

$$I_w(\theta, \lambda) = \tfrac{1}{2} \sum_e w_e(\theta_e - \lambda_e)^2,$$
$$\Gamma^*(\theta) = \max_{a \in \mathcal{I}^*(\theta)} \max_{w \in \Delta_d} \inf_{\lambda \in \mathrm{cl}(\mathrm{Alt}_a)} I_w(\theta, \lambda), \quad T^* = 1/\Gamma^*. \tag{1}$$

For a unique target Ψ, $\mathcal{I}^*(\theta) = \{\Psi(\theta)\}$; write Γ^*_Ψ, T^*_Ψ and $\mathrm{Alt}_\Psi = \mathrm{Alt}_{\Psi(\theta)}$. For matching rounds, the same game uses $x \in \mathcal{B}_n$ and I_w with $w^A_{ij} = w^B_{ji} = x_{ij}$. Write Γ^{st}_M and T^{st}_M for the stable scalar rate and time, T_p for a named blocker's time, and T^{un}_M for the unstable scalar time. The matching rate and its reciprocal are $\Gamma^{\mathrm{mat}}_M, T^{\mathrm{mat}}_M$, using the stable answer or best true blocker according to the instance; T_{mat} abbreviates the stable matching time. Standard information lower bounds (Kaufmann et al., 2016; Degenne & Koolen, 2019) give

$$\liminf_{\delta \downarrow 0} \frac{\mathbb{E}_\theta \tau_\delta}{\log(1/\delta)} \geq T^*(\theta) \qquad \text{for every } \delta\text{-correct auditor} \tag{2}$$

Appendix A records the argument. Here τ_δ counts queries or rounds according to the model, τ_{guard} denotes a guard stopping time, and $\mathrm{kl}(a, b) = a \log(a/b) + (1 - a) \log((1 - a)/(1 - b))$ is binary relative entropy, with the usual endpoint limits. This is Bernoulli Kullback–Leibler (KL) divergence.

For $a, b \geq 0$, set $h(a, b) = ab/(a + b)$ with $h(0, 0) = 0$; define the candidate information functions

$$I_p(w) = \tfrac{1}{2}(\Delta^A_p)^2 h(w^A_{ij}, w^A_{iM(i)}) + \tfrac{1}{2}(\Delta^B_p)^2 h(w^B_{ji}, w^B_{jM^{-1}(j)}), \tag{3}$$
$$J_p(x) = \tfrac{1}{2}(\Delta^A_p)^2 h(x_p, x_{e_A}) + \tfrac{1}{2}(\Delta^B_p)^2 h(x_p, x_{e_B}). \tag{4}$$

We write $I_p(w; \theta)$ and $J_p(x; \theta)$ for dependence on utilities. The notation $J_p(x; \mathsf{d})$ instead substitutes a nonnegative deficit array d for Δ. The generalized log-likelihood-ratio (GLR) statistic is $Z_a(t) = \inf_{\lambda \in \mathrm{cl}(\mathrm{Alt}_a)} \frac{1}{2} \sum_e N_e(t)(\widehat{\theta}_e(t) - \lambda_e)^2$; Z_{st} and Z_p refer to the stable and named-blocker answers. The full quadratic mean error is $L_t(\theta) = \frac{1}{2} \sum_e N_e(t)(\widehat{\theta}_e(t) - \theta_e)^2$. Empirical blockers form $\widehat{\texttt{block}}_t$. A design oracle optimizes an information game; a plug-in design uses empirical utilities. Forced exploration mixes a design with uniform sampling. Sticky selection chooses the lexicographically first pair (ordered by i, then j) within $\xi_t = t^{-1/8}$ of the best empirical rate. C-tracking samples the coordinate with largest cumulative target count minus actual count, breaking ties by coordinate index.

At scalar epoch $r = 0, 1, \ldots$, counts are $n_r = 2^r$ and the radius is $b_r = \sqrt{2 \log(\pi^2 d(r + 1)^2/(3\delta))/2^r}$. For failure level $0 < \eta < 1$ and stream count $k \geq 1$, use $b_\eta(k) = \sqrt{2 \log(\pi^2 d k^2/(3\eta))/k}$; $b = b_\delta$. Set $r_p = \min\{r : 4b_r < d_p\}$, $L_p = \log(d/\delta) + \log(1 + \log_+(1/d_p))$, $\log_+(z) = \max\{0, \log z\}$, and $L = \log(8n^2/\delta) + \log(1 + d_*^{-2})$. The stream intervals are count-uniform: jointly valid over all streams and counts. For a residual integer edge demand D, let $\ell(D)$ be its maximum row or column sum. The stopping boundaries are

$$\beta_{\mathrm{loc}}(t, \delta) = \lambda_\delta^{-1}\big[\log(P/\delta) + 2 \log L_\delta + 5 \log(t + 1) + \log(\pi^2/6)\big], \tag{5}$$
$$\beta_{\mathrm{sc}}(t, \delta) = (1 + \eta_\delta)\big[\log(1/\delta) + \tfrac{d}{2} \log(1 + t/(d\eta_\delta))\big], \tag{6}$$

where $L_\delta = \max\{2, \log(2/\delta)\}$, $\lambda_\delta = 1 - 1/L_\delta$, and $\eta_\delta = 1/\log(e/\delta)$. The dense matching boundary β_{den} replaces $\log(P/\delta), 2, 5$ in (5) by $\log(1/\delta), d/2, d + 2$.

For design certification, the simultaneous confidence event is $g^s_p \in [\ell^s_p, u^s_p]$ for every p, s; set $d^{s,-}_p = [-u^s_p]_+$, $d^{s,+}_p = [-\ell^s_p]_+$. For entrywise positive $x \in \mathcal{B}_n$ and $\lambda \in \Delta(\mathcal{P}_M)$, write

$$L^- = \min_p J_p(x; d^-), \qquad G^+ = \sum_p \lambda_p \nabla J_p(x; d^+), \qquad U^+ = \max_\pi \sum_i G^+_{i,\pi(i)}.$$

The gradient is with respect to x, and the maximum is a maximum-weight assignment over permutations. Its dual uses unrestricted row and column potentials α_i, ν_j. With exact deficits, write $L(x), G, U(x, \lambda)$; the relative tangent gap is $U/L - 1$ for $L > 0$. Design efficiency is $\min_p J_p(x; \Delta)/\Gamma_M^{\mathrm{mat}}$.

2.3 Protection structure and supplementary quantities

Definition 2.2 (Protection and dual loads)**.** The protection hypergraph has utility-coordinate vertices and one hyperedge per candidate, containing both coordinates of each rejecting comparison. Its comparison components are connected components containing a hyperedge. A pair is singly or doubly protected when one or both gains are negative. Each agent's current-partner coordinate is a star center; its protecting off-match coordinates are leaves. A least-favourable blocker law minimizes $\max_{w \in \Delta_d} \sum_p q_p I_p(w)$ over $q \in \Delta(\mathcal{P}_M)$. Let $\mathcal{R}$ index nonempty stars. In star r, leaf l belongs to candidate $p(r, l)$ and has positive deficit Δ_{rl}. Set $\rho_{rl}(q) = q_{p(r,l)}$, $\rho_r = (\rho_{rl})_l$, $b_{rl} = \rho_{rl}\Delta_{rl}^2/2$, and

$$\Lambda_r(q) = \max_{c + \sum_l x_l = 1} \sum_l b_{rl} h(c, x_l).$$

Here c is center mass and x_l leaf mass, all nonnegative. For any unit optimizer $(c_r(q), x_r(q))$, its pair contribution is

$$g_{rp}(q) = \sum_{l : p(r,l) = p} \tfrac{1}{2}\Delta_{rl}^2 h(c_r(q), x_{rl}(q)),$$

with an empty sum equal to zero; $x_r = (x_{rl})_l$.

Definition 2.3 (Pressure equilibrium)**.** A blocker-pressure equilibrium at level γ consists of distributions (q, μ) over pairs and stars such that $\mu_r > 0$ only when $\Lambda_r(q) = \gamma$, and $G_p = \sum_r \mu_r g_{rp}(q) \geq \gamma$ for all p, with equality on $\mathrm{supp}(q)$.

For a star with k leaves, suppress r and write $\Delta_l = \Delta_{rl} > 0$, $a_l = 2/\Delta_l^2$ and denote its unit-budget characteristic time by $H(a)$, also written $H(\Delta_1, \ldots, \Delta_k)$. Let z^* minimize $z + \sum_l a_l z/(z - a_l)$ over $z > \max_l a_l$. For differential formulas set $t_l = z^* - a_l$ and $S = \sum_l a_l^2/t_l^3$. A protection component's unit-budget time is T_r; for a star it is H_r. Here H is the inverse optimal minimum leaf-information rate. Vectors a majorize b when their decreasing rearrangements have larger or equal prefix sums and equal total sum.

Definition 2.4 (Structural summaries)**.** For a star, let $A_\star = \sum_l a_l$, $a_{\max} = \max_l a_l$, $k_{\mathrm{eff}} = A_\star/a_{\max}$ (effective degree), $R_{\mathrm{reuse}} = 4A_\star/H$ (center-reuse gain), and $\ell_l = \partial \log H/\partial \log a_l$ (blocker leverage). An orientation σ retains one rejecting endpoint per pair. Write $I_{p,\sigma}$ for that endpoint's term in I_p, $\Gamma_{M,\sigma} = \max_{w \in \Delta_d} \min_p I_{p,\sigma}(w)$, $\Gamma_M^{\mathrm{or}} = \max_\sigma \Gamma_{M,\sigma}$, and $R_M = \Gamma_M^{\mathrm{st}}/\Gamma_M^{\mathrm{or}}$. The margin-normalized dimensions are $D_M^{\mathrm{st}} = \kappa^2_{\Psi_M^{\mathrm{aud}}} T_M^{\mathrm{st}}$ and $D_p^{\mathrm{un}} = (\min\{g_p^A, g_p^B\}/2)^2 T_p$ for a true blocker; D_n^{st} is the supremum of D_M^{st} over stable $n \times n$ audits.

In the appendix's pilot comparison, $\widehat{w}$ is the fitted stable design or the fitted best blocker's design. It uses a lower bound $g_{\min}$ on all absolute gains, m observations per coordinate, failure levels α, ζ, relative errors $\epsilon, \eta \in (0, 1)$, exploration mass γ, $w^\gamma = (1 - \gamma)\widehat{w} + \gamma \mathbf{1}/d$, and $c_{\epsilon,\gamma} = (1 - \gamma)[(1 - \epsilon)/(1 + \epsilon)]^2$. There Γ_θ is the stable or unstable scalar audit rate.

3 Exact audit information games

3.1 Scalar queries and the distance to instability

Theorem 3.1 (Scalar audit game)**.** *On a stable instance,*

$$\Gamma_M^{\mathrm{st}} = \max_{w \in \Delta_d} \min_{p \in \mathcal{P}_M} I_p(w) = \min_{q \in \Delta(\mathcal{P}_M)} \max_{w \in \Delta_d} \sum_p q_p I_p(w).$$

This concave design has an exact second-order cone formulation with $O(n^2)$ variables and constraints. For a true blocker with gains $a, b > 0$, $T_p = 8(a^{-2} + b^{-2})$: each A-side coordinate receives mass $b^2/[2(a^2 + b^2)]$, and each B-side coordinate receives $a^2/[2(a^2 + b^2)]$. On an unstable instance, $T_M^{\mathrm{un}} = \min_{p \in \mathtt{block}(M)} T_p$.

Why a sum for stability and a minimum for a witness? To make a nonblocking pair block, an alternative must overcome *every* rejecting endpoint. The two Gaussian projection costs add, giving I_p. To destroy an existing blocker, changing *either* positive gain suffices, so its invalidation cost is the smaller endpoint cost. A stable certificate must protect against all pairs; an unstable auditor may choose its easiest true witness. Appendix A proves both cases together.

The discriminant measures the distance to an answer-changing market. Its audit margin depends on each pair's strongest rejection, so a redundant near tie need not make the audit fragile.

Theorem 3.2 (Exact stable-audit margin)**.** *If M is stable, then* $\kappa_{\Psi_M^{\mathrm{aud}}}(\theta) = \frac{1}{2}\min_{p\in\mathcal{P}_M}\max\{\Delta_p^A,\Delta_p^B\}$. *Consequently,* $\kappa_{\Psi_M^{\mathrm{aud}}}(\theta)^2/d \le \Gamma_M^{\mathrm{st}} \le \kappa_{\Psi_M^{\mathrm{aud}}}(\theta)^2/2$.

Margin intuition. An ℓ_∞ perturbation of size r changes a comparison gain by at most $2r$. Creating a blocker must overcome its strongest rejection, giving the factor $1/2$. Uniform sampling gives the lower rate bound; the general margin bound gives the upper bound. Appendix H proves both inequalities.

These bounds differ by a factor of $d/2$: the exact cost also depends on competition and sharing among comparisons.

3.2 How shared observations change the allocation

The dual law q identifies pairs that compete for samples. Every agent compares several possible partners against the same current partner. Consequently, an observation at a star's center improves several comparisons at once. The following representation separates this reuse within an agent from competition across agents.

Theorem 3.3 (Blocker-load representation)**.** *Every stable audit satisfies the exact blocker-load dual* $\Gamma_M^{\mathrm{st}} = \min_{q\in\Delta(\mathcal{P}_M)}\max_r \Lambda_r(q)$. *For fixed q, $\Lambda_r(q)$ is the maximum expected information per unit sampling mass in star r. It is convex and positively homogeneous in the blocker loads $\rho_r(q)$. When the load is nonzero, $s_r > 0$ is the unique solution of $s_r^2 = \sum_l(\sqrt{b_{rl}} - s_r)_+^2$ and $\Lambda_r = s_r^2$. The optimizing masses are $c_r = [1 + \sum_l(\sqrt{b_{rl}}/s_r - 1)_+]^{-1}$ and $x_{rl} = c_r(\sqrt{b_{rl}}/s_r - 1)_+$. Zero loads contribute zero.*

Reading the dual. For fixed q, a sampling budget spent in star r earns information at rate $\Lambda_r(q)$. The sampler therefore chooses a star with the largest rate, and the least-favourable law redistributes its weight to make this best response as weak as possible. The scalar equation in the theorem determines the active leaves: only those with $\sqrt{b_{rl}} > s_r$ receive mass. A doubly protected pair loads two stars because either endpoint can help rule it out. This is why assigning each candidate to one rejecting endpoint in advance can lose information. Appendix F proves the representation and gives the associated pressure optimality certificate.

Theorem 3.4 (Composition of disjoint comparisons)**.** *If the nonempty comparison components of the protection hypergraph are $C_1,\dots,C_m$, and T_r is the unit-budget characteristic time of auditing all blocker edges in C_r, then $T_M^{\mathrm{st}} = \sum_r T_r$. Every optimal sampling allocation assigns component r total mass $T_r/\sum_s T_s$, and every least-favourable blocker distribution assigns that component the same total probability.*

Why times add. If component r receives total mass m_r, homogeneity makes its best rate m_r/T_r. The global rate is the smallest of these component rates. Equalizing them under $\sum_r m_r = 1$ gives $m_r = T_r/\sum_s T_s$ and rate $1/\sum_s T_s$. Harder components therefore receive proportionally more samples; their disjointness does not remove competition for the same query budget. The dual masses obey the same balancing rule, as proved in Appendix F.

For singly protected audits, these components are stars. With k equal deficits Δ, a star costs $2(1+\sqrt{k})^2/\Delta^2$ and assigns center mass $1/(1+\sqrt{k})$. For example, four leaves receive mass $1/6$ each and the shared center receives $1/3$, giving time $18/\Delta^2$. Treating the four comparisons separately, with an independent center stream for each, would cost $32/\Delta^2$. Reuse saves samples, although the center still deserves more observations than any individual leaf.

3.3 Feasible deployments change the best witness

Theorem 3.5 (Matching-feasible information)**.** *With bilateral matching feedback, the stable rate is* $\Gamma_M^{\mathrm{mat}} = \max_{x \in \mathcal{B}_n} \min_p J_p(x)$*. A true blocker with gains* a, b *has optimal rate* $\min\{a^2, b^2\}/8$*, attained by an equal mixture of* M *and the matching swapping the two involved agents' partners. Thus the unstable time is* $8/\max_{p \in \mathtt{block}(M)} \min\{(g_p^A)^2, (g_p^B)^2\}$*. Every* δ*-correct auditor has asymptotic expected round count at least the applicable time times* $\log(1/\delta)$*.*

Feasibility gives $x_p + x_{e_A} \le 1$ and $x_p + x_{e_B} \le 1$, making each harmonic weight at most $1/4$. The swap mixture reaches both bounds. Consequently, the easiest witness now maximizes its smaller gain; scalar sampling instead minimizes the sum of inverse squared gains. The proof and scheduler are in Appendix B.

For a concrete comparison, suppose two true blockers have gains $(1, 1)$ and $(4/5, 4)$. Their scalar times are 16 and 13, so scalar queries prefer the second. Their matching times are 8 and $25/2$, so deployments prefer the first. Scalar queries can spend most of their budget on the weaker side of the second blocker. Matching feedback ties the two sides' exposures together and limits how often the candidate and partner edges can be observed. The same market can therefore have different easiest witnesses under the two observation models.

Table 1: Exact characteristic times for homogeneous examples. Stable: every candidate is rejected only by its A endpoint, with margin Δ. Unstable: one blocker has gains (Δ, Δ). Times multiply $\log(1/\delta)$ asymptotically.

Resource	Stable audit	One blocker
Scalar queries	$2n(1+\sqrt{n-1})^2/\Delta^2$	$16/\Delta^2$
Matching rounds	$2(1+\sqrt{n-1})^2/\Delta^2$	$8/\Delta^2$

Table 1 isolates the universal-versus-witness cost: its ratio grows as n^2 for scalar queries and as n for rounds. One round collects $2n$ utilities, so these columns measure different resources. Appendix G derives these examples and explains their margin-normalized interpretation.

4 Guards and information tracking

Guard intuition. Maintain unresolved candidate pairs. At epoch r, obtain at least 2^r observations on each candidate edge and its two partner edges. Remove a pair when one gain's upper confidence endpoint is negative; report a blocker when both lower endpoints are positive. Store a removal certificate permanently. If all pairs are removed, return stable. This rule implements the $(-1, -\varepsilon)$ example directly: it waits for a sufficient rejection, not for both signs.

For matching observations, complete the residual demand matrix D to a regular bipartite multigraph and decompose it into $\ell(D)$ perfect matchings. Keep the observations on completion edges for future epochs. This uses the fewest rounds meeting the current demands. Gain radii add the two stream radii $b(N)$ at their actual counts. Algorithm 2 and the scalar version, Algorithm 3, appear in the appendices.

Theorem 4.1 (Finite expected guard cost)**.** *The matching guard is* δ*-correct and stops almost surely on the strict audit domain. For* $0 < \delta \le 1/4$*, every stable instance satisfies*

$$\mathbb{E}_\theta \tau_{\mathrm{guard}} \le C(n + T_{\mathrm{mat}}) \log(2n) L$$

for a universal C*, and the same bound holds with probability at least* $1 - \delta$*. Every* δ*-correct matching auditor satisfies*

$$\mathbb{E}_\theta \tau \ge \max\{(n-1)(1-2\delta),\ T_{\mathrm{mat}}\, \mathrm{kl}(1-\delta, \delta)\}.$$

The cost depends on the strongest rejection d_p, avoiding inverse-square penalties for redundant near ties. The guarantee is within logarithmic factors of the lower bound; the separate tracker attains its leading constant.

Connecting the guard to the information bound. For every feasible exposure, $J_p(x) \le d_p^2 x_p$. An optimal design must therefore satisfy $x_p \ge \Gamma_M^{\mathrm{mat}}/d_p^2$ for every candidate. Summing within a row or column bounds its inverse-squared-rejection load by T_{mat}. The guard requests observations at those same scales, up to the confidence factor L. Its scheduler pays for the largest row or column load, rather than the sum of all edge demands; summing the dyadic epochs adds the $\log(2n)$ factor. A pair that remains unresolved past its nominal epoch requires a large Gaussian error on its strongest rejecting comparison. These survival probabilities decay fast enough to bound the expected extra rounds, including failed-confidence paths.

The additional n term has an operational meaning. Before certifying stability, an auditor must inspect nearly all off-matching edges in a row: an unobserved edge could hide a blocker. Since a deployment observes only one edge of that row, even large rejection gaps do not eliminate this startup cost. This accounts for the separate lower-bound term in the theorem.

Algorithm 1 Matching information tracking

Deploy n cyclic matchings to observe every edge once
for $t = n, n+1, \ldots$ **do**
  Compute empirical gains and the stable and empirical-blocker statistics
  if one statistic exceeds $\beta_{\mathrm{loc}}(t, \delta)$ **then**
    return its answer
  end if
  If empirically stable, compute a design within $t^{-1/4}$ of the optimal rate
  Otherwise select the lexicographically first blocker within ξ_t of the best estimated rate and use its swap design
  Mix the design with uniform exposure of weight $t^{-1/4}$
  Deploy one matching sampled from its Birkhoff decomposition
end for

Theorem 4.2 (Local stopping and optimal tracking)**.** *Under arbitrary adaptive matching sampling, the boundary β_{loc} controls the probability of any incorrect stable or named-blocker declaration by δ. Algorithm 1 stops almost surely and, for every fixed strict audit instance,*

$$\lim_{\delta \downarrow 0} \mathbb{E}_\theta \tau_\delta / \log(1/\delta) = T_M^{\mathrm{mat}}.$$

Any deterministic additive-error tolerance tending to zero suffices; certified relative tolerances tending deterministically to zero also suffice. The boundary depends logarithmically on n, through P, instead of carrying the dense boundary's d-dimensional penalty.

Why local evidence suffices. A wrong answer has a local witness: an overlooked blocker or a reported pair with a nonpositive gain. Each witness uses four independent utility streams but only three physical counts, since the candidate edge returns both utilities. The $5\log(t+1)$ penalty pays for at most $(t+1)^3$ count triples and a summable error budget over time; $\log P$ pays for witness selection. Thus adaptive sampling is covered without a full-market error penalty. The leading term remains $\log(1/\delta)$ at logarithmic stopping times (Appendix D).

When several blockers have equal optimal rates, choosing the empirical maximizer can keep switching designs. Sticky selection groups nearly optimal witnesses and uses a fixed ordering. Forced exploration eventually makes estimation errors smaller than its shrinking tolerance, so one optimal witness is selected consistently. This stabilization yields the claimed information rate; validity of the stopping boundary does not depend on it. Scalar tracking is in Appendix D.3.

5 Certifying a selected design

A solver's objective value describes its input means. To assess the unknown market, both estimation error and optimization error must be controlled. The following result handles them in one statement, with exact means as a special case.

Theorem 5.1 (Simultaneous information certificate)**.** *On the simultaneous gain-confidence event, every data-dependent positive $x \in \mathcal{B}_n$ and pair distribution λ with $L^- > 0$ certify that M is stable*

and

$$L^- \le \min_p J_p(x;\Delta) \le \Gamma_M^{\mathrm{mat}}(\theta) \le U^+, \qquad \textit{design efficiency} \ge L^-/U^+.$$

Stream intervals with radii $b_\eta(k)$ make this valid jointly over adaptive designs and times with probability at least $1-\eta$. With exact deficits the same result gives $L(x) \le \Gamma_M^{\mathrm{mat}} \le U(x,\lambda)$. For fixed x, the smallest tangent upper bound is the linear program minimizing $\sum_i \alpha_i + \sum_j \nu_j$ over $\lambda \in \Delta(\mathcal{P}_M)$ and unrestricted potentials satisfying $\alpha_i + \nu_j \ge G_{ij}^+$.

Intuition and proof idea. Pair rates increase with protection deficits, so lower confidence deficits give achievable information and upper deficits give an optimistic game. For every feasible y, concavity and homogeneity imply

$$\min_p J_p(y;d^+) \le \sum_p \lambda_p J_p(y;d^+) \le \langle G^+, y\rangle \le U^+.$$

The first step replaces the hardest pair by an average. The second bounds each rate by its tangent, whose constant term vanishes by homogeneity. The last step maximizes a linear function over feasible matchings. These inequalities hold for every design on the same confidence event, so choosing a design after seeing data needs no additional error allowance.

A ratio $L^-/U^+ \ge 1/2$ certifies at least half the optimal rate. If $L^- = 0$, some pair has no certified rejection yet. A small positive ratio may reflect wide intervals or a poor design; with exact deficits, only optimization and tangent-bound accuracy remain. Appendix E gives the proof and numerical safeguards.

6 Evaluation and discussion

Our synthetic experiments separate three questions: robustness to redundant gaps, stopping-boundary overhead, and design certification. Repeated looks on one path share data and are not independent trials.

Do guards avoid redundant gaps? In the scalar study, reducing a redundant rejection scale from 0.1 to 0.001 leaves the $n = 32$ guard median at $549{,}168$ observations, compared with $2{,}097{,}152$ for the same-target uniform auditor. Learning all signs grows from $34{,}430{,}976$ to $376{,}279{,}400{,}448$ observations; this is a harder diagnostic target. A separate 320-run matching-round study includes redundant (one weak and one strong rejection), mixed (one rejecting side varying by pair), double (two heterogeneous rejections), and homogeneous (equal rejections on one side) protection. Table 2 shows the $n = 32$ medians against the same-target uniform auditor. The complete settings and comparisons are in Appendix J.

Table 2: Matching-round guard study, ten paired seeds per setting. Both methods audit the same supplied matching and use the same count-uniform confidence boundary.

n	Protection	Guard rounds	Uniform rounds	Guard sec.	Uniform sec.
32	redundant	30464	131072	0.0392	0.0361
32	mixed	27744	131072	0.0398	0.0367
32	double	13216	65536	0.0341	0.0346
32	homogeneous	11136	16384	0.0328	0.032

The guard reduces deployments in every displayed family, with larger savings when protection is heterogeneous or redundant. Runtime is comparable because these simulations update Gaussian sufficient statistics in batches; fewer modeled deployments need not mean proportionally fewer seconds. The homogeneous row has less allocation freedom and shows that the guard's logarithmic-factor guarantee need not match the exact information constant.

Does local stopping reduce overhead? Table 3 uses 240 paths, thirty seeds per row. Both boundaries are evaluated on each continuing path, and all paths stop correctly without censoring.

Stable allocations use a checked numerical solver and Birkhoff sampling; the tested unstable cases have a unique blocker and use the swap design. Solver tolerance and batching are implementation choices, so this finite experiment does not establish asymptotic optimality.

Table 3: Median matching rounds at $\delta = 0.05$. Local and dense stops use the same path. Seconds include optimization and continuation to the dense stop.

n	Truth	Allocation	Local stop	Dense stop	Seconds
3	stable	tracker	1411	5379	0.0262
3	stable	uniform	1539	6787	0.0194
3	unstable	tracker	835	3331	0.0096
3	unstable	uniform	1155	4739	0.0127
6	stable	tracker	2694	37254	0.3390
6	stable	uniform	4166	67398	0.2042
6	unstable	tracker	1030	14406	0.0427
6	unstable	uniform	2694	42758	0.1182

Can a chosen design be certified? Forty numerical design checks achieve relative tangent gaps $U/L-1$ below 0.0005 in floating-point arithmetic. Separately, thirty paired-prefix paths at $n = 2, 3, 4$ use eleven looks each. All thirty eventually certify efficiency at least $1/2$; all 330 realized confidence events cover the means. These are repeated observations of thirty paths. The disclosed feasibility tolerances are numerical checks, not interval-arithmetic proofs (Appendix E).

Why the target still matters. The audit margin is $\frac{1}{2}\min_p d_p$ on a stable instance, but a margin alone does not capture competition among pairs for samples. More broadly, a tie can affect one stable-market target and leave another unchanged. Explicit 3×3 profiles have $\kappa_G \le \varepsilon/2$ but $\kappa_{M^A} \ge 1/2$, giving $T^*_G/T^*_{M^A} = \Omega(\varepsilon^{-2})$. Other profiles share their stable set and stable-pair graph but have different proposer-optimal matchings (Appendix H). Thus the requested output and the observation model must both precede a complexity claim.

Scope. The guard guarantee is finite-confidence and applies to stable audits; tracking attains the leading constant for a fixed strict instance as $\delta \downarrow 0$. Numerical certificates quantify a proposed design's rate, not its full sequential stopping cost. These distinctions explain what each theorem and experiment establishes.

Reproducibility Statement

Complete proofs, algorithms, experimental settings, seeds, and numerical tolerances are in the appendices and accompanying scripts. Appendix L maps each experiment to its code and data. Reported experiments use exact Gaussian sufficient statistics and retain the physical observation counts.

AI Use Statement

Generative AI assisted with literature search, writing, code, synthetic experiments, and derivation review. Reported numbers come from the supplied scripts. The authors are responsible for independently checking all citations, proofs, code, and outputs.

REFERENCES

Atila Abdulkadiroglu and Tayfun Sonmez. School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747, 2003. doi: 10.1257/000282803322157061.

Andreas Athanasopoulos, Anne-Marie George, and Christos Dimitrakakis. Probably correct optimal stable matching under two-sided uncertainty. In *Proceedings of the 42nd Conference on Uncertainty in Artificial Intelligence*, volume 337 of *Proceedings of Machine Learning Research*, pp. 228–254, 2026.

Jean-Yves Audibert and Sebastien Bubeck. Best arm identification in multi-armed bandits. In *Proceedings of the 23rd Conference on Learning Theory*, pp. 41–53, 2010.

Haris Aziz, Peter Biro, Serge Gaspers, Ronald de Haan, Nicholas Mattei, and Baharak Rastegari. Stable matching with uncertain linear preferences. *Algorithmica*, 82:1410–1433, 2020. doi: 10.1007/s00453-019-00650-0.

Soumya Basu. Competing bandits in matching markets via super stability. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 3226–3250, 2025.

Soumya Basu, Abishek Sankararaman, and Karthik Sankararaman. Beyond log-squared regret for decentralized bandits in matching markets. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 705–715, 2021.

Mark Braverman and Elchanan Mossel. Noisy sorting without resampling. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 268–276, 2008.

Sebastien Bubeck, Tengyao Wang, and Nitin Viswanathan. Multiple identifications in multi-armed bandits. In *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pp. 258–265, 2013.

Shicong Cen and Devavrat Shah. Regret, stability, and fairness in matching markets with bandit learners. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pp. 893–901, 2021.

Lijie Chen, Anupam Gupta, Jian Li, Mingda Qiao, and Ruosong Wang. Nearly optimal sampling algorithms for combinatorial pure exploration. In *Proceedings of the 30th Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pp. 482–534, 2017.

Shouyuan Chen, Tian Lin, Irwin King, Michael R. Lyu, and Wei Chen. Combinatorial pure exploration of multi-armed bandits. In *Advances in Neural Information Processing Systems 27*, pp. 379–387, 2014.

Remi Degenne and Wouter M. Koolen. Pure exploration with multiple correct answers. In *Advances in Neural Information Processing Systems 32*, pp. 14564–14573, 2019.

Joanna Drummond and Craig Boutilier. Elicitation and approximately stable matching with partial preferences. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, pp. 97–105, 2013.

Eyal Even-Dar, Shie Mannor, and Yishay Mansour. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal of Machine Learning Research*, 7:1079–1105, 2006.

Tomás Feder. A new fixed point approach for stable networks and stable marriages. *Journal of Computer and System Sciences*, 45(2):233–284, 1992. doi: 10.1016/0022-0000(92)90048-N.

Uriel Feige, Prabhakar Raghavan, David Peleg, and Eli Upfal. Computing with noisy information. *SIAM Journal on Computing*, 23(5):1001–1018, 1994. doi: 10.1137/S0097539791195877.

Tanner Fiez, Lalit Jain, Kevin G. Jamieson, and Lillian J. Ratliff. Sequential experimental design for transductive linear bandits. In *Advances in Neural Information Processing Systems 32*, pp. 10667–10677, 2019.

David Gale and Lloyd S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15, 1962. doi: 10.2307/2312726.

Aurelien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In *Proceedings of the 29th Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pp. 998–1027, 2016.

Begum Genc, Mohamed Siala, Barry O'Sullivan, and Gilles Simonin. Robust stable marriage. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp. 4925–4926, 2017. doi: 10.1609/aaai.v31i1.11107.

Dan Gusfield. Three fast algorithms for four problems in stable marriage. *SIAM Journal on Computing*, 16(1):111–128, 1987. doi: 10.1137/0216010.

Dan Gusfield and Robert W. Irving. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press, Cambridge, MA, 1989.

John William Hatfield and Paul R. Milgrom. Matching with contracts. *American Economic Review*, 95(4):913–935, 2005. doi: 10.1257/0002828054825466.

Reinhard Heckel, Nihar B. Shah, Kannan Ramchandran, and Martin J. Wainwright. Active ranking from pairwise comparisons and when parametric assumptions do not help. *Journal of Machine Learning Research*, 20(126):1–48, 2019.

Hadi Hosseini, Sanjukta Roy, and Duohan Zhang. Putting gale and shapley to work: Guaranteeing stability through learning. In *Advances in Neural Information Processing Systems 37*, 2024.

Nicole Immorlica and Mohammad Mahdian. Marriage, honesty, and stability. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 53–62, 2005.

Robert W. Irving. An efficient algorithm for the stable roommates problem. *Journal of Algorithms*, 6 (4):577–595, 1985. doi: 10.1016/0196-6774(85)90033-1.

Robert W. Irving. Stable marriage and indifference. *Discrete Applied Mathematics*, 48(3):261–272, 1994. doi: 10.1016/0166-218X(92)00179-P.

Robert W. Irving and Paul Leather. The complexity of counting stable marriages. *SIAM Journal on Computing*, 15(3):655–667, 1986. doi: 10.1137/0215048.

Meena Jagadeesan, Alexander Wei, Yixin Wang, Michael I. Jordan, and Jacob Steinhardt. Learning equilibria in matching markets from bandit feedback. *Journal of the ACM*, 70(3):1–46, 2023.

Kevin G. Jamieson and Robert D. Nowak. Active ranking using pairwise comparisons. In *Advances in Neural Information Processing Systems 24*, pp. 2240–2248, 2011.

Kevin G. Jamieson, Matthew Malloy, Robert Nowak, and Sebastien Bubeck. lil'ucb: An optimal exploration algorithm for multi-armed bandits. In *Proceedings of the 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pp. 423–439, 2014.

Marc Jourdan, Mojmír Mutný, Johannes Kirschner, and Andreas Krause. Efficient pure exploration for combinatorial bandits with semi-bandit feedback. In *Algorithmic Learning Theory*, volume 132 of *Proceedings of Machine Learning Research*, pp. 805–849. PMLR, 2021.

Shivaram Kalyanakrishnan, Ambuj Tewari, Peter Auer, and Peter Stone. Pac subset selection in stochastic multi-armed bandits. In *Proceedings of the 29th International Conference on Machine Learning*, pp. 655–662, 2012.

Zohar S. Karnin, Tomer Koren, and Oren Somekh. Almost optimal exploration in multi-armed bandits. In *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pp. 1238–1246, 2013.

Emilie Kaufmann and Wouter M. Koolen. Mixture martingales revisited with applications to sequential tests and confidence intervals. *Journal of Machine Learning Research*, 22(246):1–44, 2021.

Emilie Kaufmann, Olivier Cappe, and Aurelien Garivier. On the complexity of best-arm identification in multi-armed bandit models. *Journal of Machine Learning Research*, 17(1):1–42, 2016.

Donald E. Knuth. *Mariages Stables*. Les Presses de l'Universite de Montreal, Montreal, 1976.

Fang Kong and Shuai Li. Player-optimal matching learning in the bandit setting. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 3504–3523, 2023.

Fang Kong, Yichi Yin, and Shuai Li. Improved analysis for bandit learning in matching markets. In *Advances in Neural Information Processing Systems 37*, pp. 91904–91929, 2024.

Fang Kong, Yichi Yin, and Shuai Li. Bandit learning in matching markets with indifference. In *Proceedings of the Thirteenth International Conference on Learning Representations*, 2025.

Lydia T. Liu, Horia Mania, and Michael I. Jordan. Competing bandits in matching markets. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 1618–1628, 2020.

Lydia T. Liu, Feng Ruan, Horia Mania, and Michael I. Jordan. Bandit learning in decentralized matching markets. *Journal of Machine Learning Research*, 22(211):1–34, 2021.

David F. Manlove, Robert W. Irving, Kazuo Iwama, Shuichi Miyazaki, and Yasufumi Morita. Hard variants of stable marriage. *Theoretical Computer Science*, 276(1–2):261–279, 2002. doi: 10.1016/S0304-3975(01)00206-4.

Tejas Pagare, Agniv Bandyopadhyay, and Sandeep Juneja. Optimal algorithms for bandit learning in matching markets. *arXiv preprint arXiv:2509.14466*, 2025.

Riccardo Poiani, Martino Bernasconi, and Andrea Celli. Non-asymptotic analysis of (sticky) track-and-stop. In *Proceedings of the Fourteenth International Conference on Learning Representations*, 2026.

Baharak Rastegari, Anne Condon, Nicole Immorlica, and Kevin Leyton-Brown. Two-sided matching with partial information. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, pp. 152–158, 2013.

Alvin E. Roth. The evolution of the labor market for medical interns and residents: A case study in game theory. *American Economic Review*, 74(6):991–1016, 1984.

Uriel G. Rothblum. Characterization of stable matchings as extreme points of a polytope. *Mathematical Programming*, 54:57–67, 1992. doi: 10.1007/BF01586041.

Abishek Sankararaman, Soumya Basu, and Karthik Sankararaman. Dominate or delete: Decentralized competing bandits in serial dictatorship. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pp. 1252–1260, 2021.

Chung-Piaw Teo and Jay Sethuraman. The geometry of fractional stable matchings and its applications. *Mathematics of Operations Research*, 23(4):874–891, 1998. doi: 10.1287/moor.23.4.874.

John H. Vande Vate. Linear programming brings marital bliss. *Operations Research Letters*, 8(3): 147–153, 1989. doi: 10.1016/0167-6377(89)90007-2.

Zheshun Wu, Jinhang Zuo, Zenglin Xu, and Fang Kong. Bandit learning in matching markets robust to adversarial corruptions. In *Proceedings of the Fourteenth International Conference on Learning Representations*, 2026.

A PROOFS OF THE SCALAR INFORMATION GAME

A.1 INFORMATION LOWER BOUND

The following change-of-measure argument explains the benchmark (2); it is standard rather than an additional audit theorem.

Lemma A.1 (Stopped change of measure)**.** *Let an adaptive algorithm sample independent Gaussian coordinates with unit variance and stop at an almost surely finite stopping time τ. For parameters θ, λ and every event E measurable at τ,*

$$\frac{1}{2}\sum_{e=1}^{d} \mathbb{E}_\theta[N_e(\tau)](\theta_e - \lambda_e)^2 \geq \mathrm{kl}\big(\mathbb{P}_\theta(E), \mathbb{P}_\lambda(E)\big).$$

Proof. Let L_τ be the log likelihood ratio of the adaptively collected observations under θ against λ. Conditional on the sampled coordinate at each round, the expected one-step log likelihood ratio under θ is $(\theta_e - \lambda_e)^2/2$. Optional sampling applied to the compensated likelihood process, followed by monotone convergence through bounded stopping times, gives $\mathbb{E}_\theta L_\tau = \frac{1}{2}\sum_e \mathbb{E}_\theta[N_e(\tau)](\theta_e - \lambda_e)^2$. The data-processing inequality for relative entropy under the measurable map that records only $\mathbf{1}\{E\}$ yields $\mathbb{E}_\theta L_\tau \geq \mathrm{kl}(\mathbb{P}_\theta(E), \mathbb{P}_\lambda(E))$. □

Justification of (2). For a unique correct answer a, first fix $\lambda \in \mathrm{Alt}_a$ and let E be the event that the algorithm outputs a. By δ-correctness, $\mathbb{P}_\theta(E) \geq 1 - \delta$ and $\mathbb{P}_\lambda(E) \leq \delta$. Applying Lemma A.1, dividing the expected counts by $\mathbb{E}_\theta\tau$, and taking the infimum over $\lambda \in \mathrm{Alt}_a$ gives the usual answer-specific information constraint. For every fixed allocation the quadratic information is continuous in λ, so its infimum over Alt_a equals its infimum over $\mathrm{cl}(\mathrm{Alt}_a)$; no correctness claim is needed at a boundary parameter where more than one answer may be valid. Maximizing over allocations therefore gives $\mathbb{E}_\theta\tau \geq T^*(\theta)\,\mathrm{kl}(1-\delta, \delta)$. Since $\mathrm{kl}(1-\delta,\delta)/\log(1/\delta) \to 1$, the asymptotic claim follows.

For finitely many correct answers, the preceding binary contraction need not be sharp because the algorithm may randomize among answers. The multiple-answer lower-bound theorem of Degenne & Koolen (2019) applies to independent one-parameter exponential-family coordinates and to any finite correct-answer correspondence. Gaussian coordinates satisfy its assumptions, and their coordinate-wise divergence is $(\theta_e - \lambda_e)^2/2$. Substitution yields exactly the outer maximum over $a \in \mathcal{I}^*(\theta)$ in the displayed definition of $\Gamma^*(\theta)$ and proves the stated bound. □

A.2 COMPARISON PROJECTION AND CONCAVITY

Lemma A.2 (Gaussian comparison calculus)**.** *For $y - x = \Delta > 0$ and $r, s \geq 0$,*

$$\inf_{x' \geq y'} \tfrac{1}{2}\{r(x - x')^2 + s(y - y')^2\} = \tfrac{1}{2}h(r, s)\Delta^2.$$

The function h is continuous, concave, coordinatewise nondecreasing, and positively homogeneous on the nonnegative quadrant.

Proof. If $r = 0$ or $s = 0$, the unobserved coordinate can be moved to the boundary at zero cost. Assume $r, s > 0$. The feasible halfspace is closed and convex, and the unconstrained minimizer (x, y) is infeasible. The constrained minimizer therefore lies on $x' = y' = m$. Differentiating $r(x - m)^2 + s(y - m)^2$ gives $m = (rx + sy)/(r + s)$. Substitution yields $r[s\Delta/(r + s)]^2 + s[r\Delta/(r + s)]^2 = rs\Delta^2/(r + s)$, and multiplication by $1/2$ completes the proof.

Continuity at the origin follows from $0 \leq h(x, y) \leq \min\{x, y\}$. Positive homogeneity and monotonicity are immediate. On the positive orthant its Hessian is

$$\nabla^2 h(x, y) = -\frac{2}{(x + y)^3}\begin{pmatrix} y^2 & -xy \\ -xy & x^2 \end{pmatrix},$$

which is negative semidefinite because the matrix after the minus sign is $(y, -x)^\top(y, -x)$. Concavity extends to the boundary by continuity.

□

A.3 Stable and unstable answers

Proof of Theorem 3.1. For a candidate pair, the closure of the blocking event is the intersection of its two nonnegative-gain halfspaces. They involve disjoint utility coordinates. A nonnegative gain needs no displacement; a negative gain costs one half its squared deficit times the harmonic weight, by Lemma A.2. The two costs therefore add to $I_p(w)$. The alternative to a stable answer is the finite union of these blocking cones. Its distance is the minimum of their distances, giving the stable max–min formula.

The functions I_p are continuous and concave. Maximizing t subject to $w \in \Delta_d$ and $t \le I_p(w)$ is therefore a convex optimization problem. For a fixed w, the minimum over p equals the minimum of $\sum_p q_p I_p(w)$ over the pair simplex. This payoff is affine in q and concave in w on compact convex domains; Sion's minimax theorem exchanges the extrema. At a saddle point, positive q_p can occur only at a minimum-information pair.

For an exact conic representation, $z \le h(x, y)$ with nonnegative variables is equivalent to $x \ge z$, $y \ge z$, and $(x - z)(y - z) \ge z^2$. This last inequality is a rotated second-order cone. Introduce one such variable per protecting comparison and require t to be no larger than the sum of its squared-deficit-weighted variables for every pair. There are $O(n^2)$ variables and constraints. For rational inputs, additive accuracy in the rate is computable in polynomial time in the encoding size and $\log(1/\epsilon)$. Converting a rate error to an absolute error in its reciprocal also requires controlling the positive rate; an additive rate tolerance alone is not an absolute-time tolerance.

A named blocker becomes invalid if either positive gain is reduced to zero. Its information is therefore the minimum of its two endpoint projection costs. If side A receives total mass s, its largest harmonic weight is $s/4$, attained by splitting that mass equally. The other side similarly receives harmonic weight $(1 - s)/4$. Thus the best rate is $\max_{s\in[0,1]} \min\{sa^2/8, (1 - s)b^2/8\}$. Equalizing the terms gives the allocation and T_p in the theorem. Finally, the multiple-answer game chooses the true blocker with largest rate, equivalently smallest time. □

B Matching feedback and the demand scheduler

A matching round chooses a permutation L based on previous observations, then receives $u_{iL(i)} + Z_i^A$ and $v_{L(i)i} + Z_i^B$ for every i, with fresh independent standard Gaussian noises. The pair exposure $N_{ij}(t)$ counts rounds containing (i, j). Therefore its row and column sums both equal t. This equality holds pathwise, including for randomized adaptive deployment. For integrable stopping times, the expected counts divided by the expected number of rounds are doubly stochastic. Independent coordinates on the same deployed edge have the same exposure, but different Gaussian noise; these two constraints must not be conflated.

For the stable answer, the alternative set is the union over pairs of the cones in which both gains are positive. Its closure allows weak inequalities. The Gaussian KL cost against alternative utilities $(\widetilde{u}, \widetilde{v})$ per round is $\frac{1}{2}\sum_{ij} x_{ij}[(u_{ij} - \widetilde{u}_{ij})^2 + (v_{ji} - \widetilde{v}_{ji})^2]$. Minimizing over the two coordinates of each violated comparison gives its squared negative gain times the harmonic weight divided by two. Coordinates on opposite sides belong to different agents' utility lists and are distinct, even when their physical edges coincide. Their costs therefore add, giving $J_p(x)$. The closest cone gives the minimum over p. Maximization over the Birkhoff polytope proves the stable formula.

A fixed blocking answer p becomes incorrect if either gain is nonpositive. The corresponding information is $\min\{a^2 h(x_p, x_{e_A})/2, b^2 h(x_p, x_{e_B})/2\}$. For nonnegative s, t with $s + t \le 1$, the identity $st/(s + t) \le (s + t)/4 \le 1/4$ proves the upper bound. Let $i' = M^{-1}(j)$. In the swap matching, i receives j, i' receives $M(i)$, and all remaining assignments agree with M. The equal mixture has exposure $1/2$ on the candidate and on both original partner edges. It reaches both harmonic bounds, including at $n = 2$. Maximizing over correct blocking answers gives the stated witness-selection criterion. The multiple-answer lower bound is the established answer-specific information-game lower bound; it is not obtained by incorrectly applying a single fixed-answer event to an algorithm that may choose among several blockers.

For the homogeneous stable example, write $k = n - 1$ and fix one row's diagonal mass s. At least one leaf has mass at most $(1 - s)/k$, and h is increasing in either argument. Hence the row's

minimum harmonic weight is at most $s(1-s)/(ks+1-s)$. Equal leaf masses attain this value. Differentiation on $[0,1]$ gives the maximum $(1+\sqrt{k})^{-2}$ at $s=(1+\sqrt{k})^{-1}$. The matrix with $x_{ii}=s$ and $x_{ij}=(1-s)/k$ for $i\neq j$, at this maximizing s, attains that value in every row and column. This proves the exact optimum without assuming that an optimizer must be symmetric. If comparison gains on the other side are positive, their stable-deficit terms vanish; their observations are returned but provide no local protection information. Thus the bilateral feedback model cannot be replaced by twice as many freely chosen scalar queries.

Lemma B.1 (Minimum rounds for prescribed demands)**.** *For any nonnegative integer $n\times n$ demand matrix D, the minimum number of perfect-matching rounds whose cumulative edge multiplicities dominate D is $\ell(D)$.*

Proof. One round contributes one observation to every row and column, so all row and column demand sums are lower bounds. Put $\ell=\ell(D)$; if $\ell=0$, there is no demand and no round is needed. Row deficits $\ell-\sum_j D_{ij}$ and column deficits $\ell-\sum_i D_{ij}$ are nonnegative integers with the same total. Fill them greedily to obtain an integer matrix $C\geq D$ with all row and column sums ℓ. In its support graph, any row set S has total multiplicity $\ell|S|$ incident to its neighbor set $N(S)$. The latter can absorb at most $\ell|N(S)|$, so $|N(S)|\geq|S|$. Hall's condition gives a perfect matching. Remove one copy of it and repeat. This decomposes C into exactly ℓ perfect matchings. The implementation removes the minimum positive multiplicity along a found matching in a single batch; this only compresses the description of consecutive equal deployments. □

Algorithm 2 Matching-feasible guard sampling

1: Keep every off-matching pair active and initialize all edge counts and bilateral sums to zero
2: **for** $r=0,1,2,\ldots$ **do**
3: Mark every active candidate edge and its two current-partner edges as required
4: Set $D_{ij}=\max\{0,2^r-N_{ij}\}$ on required edges and zero elsewhere
5: Complete and decompose D as in Lemma B.1
6: Deploy every resulting matching batch; retain both utilities on every edge, including fill edges
7: Form all required gain intervals by adding the two coordinate radii $b(N)$
8: Return any pair with two strictly positive lower endpoints; remove pairs with one strictly negative upper endpoint
9: Return stable if no pair remains
10: **end for**

For a fixed coordinate and deterministic prefix length $N\geq 1$, the Gaussian bound is $\Pr\{|\widehat{\theta}_{e,N}-\theta_e|>b(N)\}\leq 6\delta/(\pi^2 dN^2)$. Summing over N and e gives failure probability at most δ. A coordinate stream can be imagined in advance; adaptive matching choices only reveal prefixes of these streams. Thus the same event covers every random sample count, and batching creates no selection bias. On this event an excluded candidate is not a blocker, and any returned blocker is valid. A stable declaration has excluded all candidates, proving correctness.

For almost-sure termination, use strong laws for all finitely many coordinate streams, rather than conditioning on the confidence event. If a pair stayed active forever, its candidate and two partner edges would receive at least 2^r observations by epoch r. The radii at these counts tend to zero. Every nonblocking pair in the strict audit domain has a strictly negative gain, and every blocking pair has two strictly positive gains. Each surviving pair is therefore eventually resolved almost surely. On an early confidence failure the algorithm could stop with a wrong answer, but it does not silently drop that trajectory from the correctness calculation. The experimental records retain decisions, batch counts, time, and scalar observations.

The demand scheduler is optimal for the specified current demands, not for the complete unknown-instance audit problem. The information program describes the latter's asymptotic lower-bound geometry. We do not infer information-optimality of this epoch scheduler from its optimal per-epoch decomposition.

Algorithm 3 Adaptive guard racing

Require: matching M, confidence $\delta \in (0, 1/2)$

$U \leftarrow \mathcal{P}_M$ (unresolved pairs); initialize coordinate streams
for $r = 0, 1, \ldots$ **do**
  sample each $e \in \bigcup_{p \in U} Q_p$ to cumulative count 2^r
  for $p \in U$ **do**
    if $\widehat{g}_p^A - 2b_r > 0$ and $\widehat{g}_p^B - 2b_r > 0$ **then**
      return blocker p
    else if $\widehat{g}_p^A + 2b_r < 0$ or $\widehat{g}_p^B + 2b_r < 0$ **then**
      remove p from U
    end if
  end for
  if $U = \varnothing$ **then**
    return `stable`
  end if
end for

C Finite-confidence guard guarantees

C.1 Scalar queries

Theorem C.1 (Scalar guard complexity)**.** *Algorithm 3 is δ-correct and stops almost surely on every strict audit instance. On a stable instance, use the quantities in Section 2. With probability at least $1 - \delta$, its total sample count obeys $\tau \le \sum_{e:\exists p, e \in Q_p} 2^{\max_{p:e \in Q_p} r_p}$. For a universal constant C_0, it also obeys $\mathbb{E}_\theta \tau \le C_0 \sum_p (1 + d_p^{-2}) L_p$. Conversely, every δ-correct audit algorithm satisfies $\mathbb{E}_\theta \tau \ge \max\{n(n-1)(1-2\delta),\ 2\,\mathrm{kl}(1-\delta, \delta) \sum_p d_p^{-2}\}$ on every stable strict instance.*

Proof of Theorem C.1. Fix independent infinite Gaussian streams for all utility coordinates before running the algorithm. Adaptive sampling only reveals prefixes. At epoch r, a used prefix has deterministic length 2^r, even though whether it is revealed depends on earlier data. A two-sided Gaussian tail gives $\Pr(|\widehat{\theta}_{e,2^r} - \theta_e| > b_r) \le 2\exp(-2^r b_r^2/2) = 6\delta/(\pi^2 d(r+1)^2)$. Summing over e and r gives total probability at most δ. On the complementary event, the difference of two coordinate means has error at most $2b_r$. The two tests in Algorithm 3 are therefore correct whenever they are used. Removed comparisons are never revisited; later changes to a shared baseline mean do not change a past certificate.

For a stable pair p, select once and for all an endpoint with true gain $g_p = -d_p$, and denote its empirical gain by $\widehat{g}_p$. Its two coordinates have equal count 2^r whenever p is unresolved at epoch r. If $4b_r < d_p$, the confidence event implies $\widehat{g}_p + 2b_r < 0$. Every incident coordinate therefore has final count at most $2^{\max_{p:e \in Q_p} r_p}$. This proves the high-probability budget without double-counting shared coordinates.

For expectation, let $A_{p,r}$ mean that p survives epoch r and no answer has yet been returned. For $r \ge r_p$, this event requires $\widehat{g}_p - g_p \ge d_p - 2b_r > d_p/2$. The prefix difference has variance $2/2^r$, so $\Pr(A_{p,r}) \le \exp(-2^r d_p^2/16)$. This implication holds unconditionally, including histories on which an earlier confidence interval failed. Charging all observations at epoch $r+1$ to pairs still active after epoch r gives $\mathbb{E}\tau \le 4\sum_p [\sum_{r=0}^{r_p} 2^r + \sum_{r=r_p+1}^{\infty} 2^r \exp(-2^{r-1} d_p^2/16)]$. For any $a > 0$, the dyadic sum $\sum_{r \ge 0} 2^r e^{-a2^r}$ is at most C/a for a universal C: compare each term with the integral on the preceding dyadic interval and handle the first interval separately. Thus the last display is at most $C \sum_p (2^{r_p} + d_p^{-2})$.

We record explicitly how to solve the epoch threshold. Write $H = \log(\pi^2 d/(3\delta))$, so the requirement is $2^r d_p^2 > 32[H + 2\log(r+1)]$. For $d_p \ge 1$, taking 2^r a sufficiently large constant multiple of H satisfies this inequality. For $d_p < 1$, take 2^r between $K d_p^{-2}[H + \log(1 + \log(1/d_p))]$ and twice that number. Its logarithm is at most a constant plus $2\log(1/d_p) + \log K + \log[H + \log(1 + \log(1/d_p))]$. Hence $\log(r+1)$ is at most a constant multiple of $H + \log(1 + \log(1/d_p))$; increasing universal K makes the required inequality strict. This proves the stated expectation bound. Gaussian tails are

summable, so every active pair's comparisons converge along epochs; a stable instance loses every pair, and on an unstable instance a surviving true blocker is eventually certified. Thus stopping is almost sure on the stated strict domain.

For the lower bound, denote the two off-match utility coordinates for pair p by $e_A(p), e_B(p)$. For $\varepsilon > 0$, increase each negative gain at that pair to ε, leaving an already positive gain and all other utility coordinates unchanged. Choose ε outside the finite set producing an unwanted tie. The resulting alternative makes p block the proposed matching. Apply the stopped-experiment change-of-measure inequality to the stable-output event, whose probability is at least $1 - \delta$ under the true stable market and at most δ under the alternative. It yields $\frac{1}{2}(d_p + \varepsilon)^2 \mathbb{E}[N_{e_A(p)} + N_{e_B(p)}] \geq \mathrm{kl}(1 - \delta, \delta)$. Let ε decrease to zero. Each off-match coordinate belongs to exactly one candidate pair, so summing proves the inverse-square lower bound. Infinite expected stopping time makes the inequality immediate and requires no integrability reduction.

For the unit-sample term, consider the event of stopping with a stable answer without ever sampling either off-match coordinate of p. Its probability is unchanged by this alternative: every revealed observation and every sampling probability on such a path have the same law. It is therefore at most δ. Correctness at the true stable market gives probability at least $1 - 2\delta$ of sampling one of those coordinates. Summing over pairs proves $\mathbb{E}\tau \geq n(n-1)(1-2\delta)$. This argument does not assume that the learner first determines which endpoint is protective. □

C.2 Deployment rounds

Proof of Theorem 4.1. Relabel so the supplied matching is diagonal. We prove Theorem 4.1 for the exact algorithm in `run_guard_rounds.py`. All counts below are physical-edge counts; one deployed edge supplies both utility coordinates. The stored certificate for a removed pair is not revoked when its current-partner edge is sampled again.

Let $H_p = 1 + d_p^{-2}$ and $L = \log(8n^2/\delta) + \log(1 + d_*^{-2})$. Fix a sufficiently large universal K, and define R_p by $2^{R_p - 1} < KH_pL \leq 2^{R_p}$. Since $\log H_p \leq L$ and $\log L \leq L$, increasing K ensures $4b(N) < d_p$ for every $N \geq 2^{R_p}$; it also ensures $2^{R_p} d_p^2 \geq KL$. The count-uniform radius is decreasing in the count on this range. Integer round-up only changes a universal factor.

For an off-matching edge p, define $R_e = R_p$. For a diagonal edge, take the largest R_p among all candidate pairs using it. Put $w_e = 2^{R_e}$ and $W = \max\{\max_i \sum_j w_{ij}, \max_j \sum_i w_{ij}\}$. Since each harmonic term is at most x_p and both rejecting gaps are at most d_p, $J_p(x) \leq d_p^2 x_p$. At a max–min optimizer x, $x_p \geq \Gamma_M^{\mathrm{mat}} / d_p^2$. Summing in each row and column gives the inverse-margin load at most T_{mat}. The weight on a diagonal is at most a constant times $L(1 + T_{\mathrm{mat}})$; the sum on its off-diagonal row or column is at most a constant times $L(n - 1 + T_{\mathrm{mat}})$. Hence $W \leq C_1 L(n + T_{\mathrm{mat}})$.

At epoch r, first charge demands of active pairs with $r \leq R_p$ to early edges with $R_e \geq r$. Each residual edge demand is at most 2^r, even when completion edges were previously oversampled. For every row or column, the corresponding early demand is at most $n2^r$ and at most its total w-weight. Thus its maximum load is at most $\min\{n2^r, W\}$. No early edges remain for $r > \max_e R_e \leq \log_2 W$. The sum of $n2^r$ up to $2^r \leq W/n$ is at most $2W$; after this point there are at most $2 + \log_2 n$ nonzero epochs, each costing at most W. If $W < n$, the latter bound covers every epoch directly. The total early schedule is at most $C_2 W \log(2n)$.

This deterministic accounting remains valid on paths with false certificates: removing a pair can only decrease subsequent demands. Late pairs may add demands. The functional giving maximum row or column load is subadditive. Each pair uses at most three physical edges, so charge at most $3 \cdot 2^r$ extra rounds to a late pair surviving until epoch r. This bounds the actual minimal-demand scheduler even if early and late edges overlap.

For a late pair p, at epoch $r - 1 \geq R_p$ all its comparison edges have at least $k = 2^{r-1}$ observations. If its strongest rejecting comparison has not been certified, one of its two utility streams has sample-mean error of absolute value at least $d_p/4$ at some count $N \geq k$. For independent standard Gaussian observations with partial sums S_N, the exponential martingale $\exp(\epsilon S_N - \epsilon^2 N/2)$ and its negative counterpart give $\Pr\{\exists N \geq k : |S_N|/N \geq \epsilon\} \leq 2e^{-k\epsilon^2/2}$. This bound handles adaptively selected and oversampled counts without conditioning on the stopping time. Union-bounding the two streams

yields $4e^{-2^{r-1}d_p^2/32}$. An early unstable answer, correct or incorrect, only ends the run and reduces this survival probability.

For $a = 2^{R_p}d_p^2/32 \geq 2$, the dyadic tail satisfies $\sum_{r>R_p} 2^r e^{-2^{r-1}d_p^2/32} \leq C_3 d_p^{-2} e^{-a/2}$. This follows by comparing consecutive terms, whose ratio is at most $2e^{-a}$, and absorbing ae^{-a} into a constant multiple of $e^{-a/2}$. Choose K so $e^{-a/2} \leq n^{-3}$. Since $\sum_p H_p \leq n(n-1) + nT_{\mathrm{mat}}$, the expected late cost is $O(n + T_{\mathrm{mat}})$. This proves the expectation bound without discarding failed-confidence trajectories. On the simultaneous count-uniform confidence event, all pairs are removed by their deadlines and the early bound alone proves the high-probability assertion.

For completeness, a Gaussian change of measure applies to the stable-output event for each strict unstable alternative. The expected exposure matrix, normalized by finite expected rounds, lies in $\mathcal{B}_n$. Minimizing the divergence over a blocking-pair alternative and maximizing over feasible exposure matrices gives the claimed nonasymptotic information lower bound. If the expectation is infinite the inequality is automatic. To obtain the startup term, modify one off-diagonal physical edge to create a strict blocker. Until it is observed, the two experiments can be coupled exactly. The probability of returning stable without observing it is at most δ under the original instance. Since stable output itself has probability at least $1 - \delta$, the edge is observed with probability at least $1 - 2\delta$. Sum these probabilities over the $n - 1$ off-diagonal edges in one row; there are at most τ distinct observed edges in that row. This proves the startup lower bound and finishes the theorem. □

D Local stopping and information tracking

D.1 Witness-local error control

Fix a candidate $p = (i, j)$ and relabel so that the proposed matching is the identity. The physical edges are $(i, j), (i, i), (j, j)$. The first edge reveals two independent utilities; each partner edge supplies the one relevant utility. For deterministic prefix lengths $a, b, c \leq t$, define $V_p(a, b, c)$ to be one half the sum of count times squared mean error over these four streams, using zero contribution at a zero count. Its nonzero standardized errors are independent standard Gaussians. Hence $\mathbb{E}\exp(\lambda V_p) \leq (1 - \lambda)^{-2}$ for $0 < \lambda < 1$. Pre-generating independent infinite streams on each physical edge justifies applying this fact to every deterministic triple before the algorithm chooses which prefixes to inspect.

For the boundary in Theorem 4.2, Markov's inequality for one triple gives probability at most $\delta/[P(t + 1)^5(\pi^2/6)]$. There are at most $(t + 1)^3$ triples for one pair at time t. Summing over pairs and all integer $t \geq 0$ bounds the total by δ. No independence between different pairs, count choices, or rounds is assumed in this union bound.

If the auditor wrongly returns stable, fix any truly blocking pair. The true four-dimensional mean vector lies in the closure of that pair's blocking cone, so the squared distance from the empirical vector to this cone is no greater than V_p. The stable statistic, a minimum over all pairs, is no larger. If the auditor wrongly returns pair p, the true four-dimensional vector lies in the complement of its blocking cone; the named-blocker projection has the same upper bound. Thus every error entails crossing of one of the local error processes. This also handles boundary parameters for correctness, although finite stopping and the asymptotic theorem retain the strict-comparison assumption.

D.2 Attainment in expected deployment rounds

Proof of the attainment claim in Theorem 4.2. We specify both a valid boundary and an allocation rule, because feasible exposure vectors alone do not prove attainability. Begin with n cyclic perfect matchings, obtaining one observation on every edge. Continue the sampling rule after any hypothetical stopping time for purposes of analysis. At time t, estimate the utility means and determine whether the proposed matching is empirically stable. If so, choose a feasible matrix w_t whose estimated stable-game value is within $t^{-1/4}$ of the optimum. If not, compute each empirical blocker's estimated round rate $\min\{(\widehat{g}_p^A)^2, (\widehat{g}_p^B)^2\}/8$. Choose the lexicographically first blocker whose estimated rate is within $t^{-1/8}$ of the best, and let w_t be its half–half swap design. Set $x_t = (1 - t^{-1/4})w_t + t^{-1/4}\mathbf{1}\mathbf{1}^\top/n$. A Birkhoff decomposition realizes x_t by sampling one perfect matching with this conditional exposure matrix. The initial finitely many rounds may be treated separately.

Use the local boundary proved above and stop when an empirically valid answer exceeds it. An incorrect declaration has probability at most δ under any adaptive sampling rule. It remains to prove termination and the optimal expected cost for this allocation rule.

We give the tail argument for expectation, not only almost-sure convergence. Each edge has conditional probability at least $t^{-1/4}/n$ of observation in round t. Bounded martingale differences imply, for large t, an edge-count lower bound $c_1 t^{3/4}$ with failure probability at most $c_2 e^{-c_3 t^{1/2}}$, where constants may depend on the fixed market size. A union bound over possible deterministic prefix lengths and Gaussian mean tails then gives $\Pr\{\|\widehat{\theta}(t) - \theta\|_\infty > t^{-1/4}\} \le A e^{-ct^{1/4}}$ after adjusting constants. All finitely many current-versus-candidate gains are nonzero, so sufficiently accurate estimates identify the stability label and the true blocker set.

On stable instances, $\min_p J_p(x;\theta)$ is concave and uniformly continuous in $x \in \mathcal{B}_n$, and uniformly Lipschitz in θ on a fixed compact neighborhood of the true parameter. The same holds for the individual blocker information functions. The harmonic mean is globally Lipschitz on the nonnegative square: its partial derivatives away from the origin are bounded by one, and continuity handles the origin. Consequently, an approximately optimal estimated design has true value $\Gamma_M^{\mathrm{mat}} - o(1)$. Forced exploration changes this by $o(1)$. On unstable instances, the estimated rates converge faster than the tolerance $t^{-1/8}$ on the above high-probability event. The tolerance includes every truly optimal witness and eventually excludes every strictly suboptimal witness. Lexicographic choice therefore stabilizes on one true optimal witness, even in the presence of ties among optimal rates. Its chosen swap design is exactly rate-optimal.

For an explicit high-probability tracking statement at time t, require the preceding estimation event to hold for all times from $\lceil \sqrt{t} \rceil$ through t. Its complement is bounded by a summable tail $A' e^{-c' t^{1/8}}$ after absorbing polynomial factors. The first $\sqrt{t}$ rounds have vanishing relative weight. Concavity shows that the average of the subsequent true near-optimal designs has information at least $\Gamma_M^{\mathrm{mat}} - o(1)$ for the stable answer or the stabilized optimal blocker. Another bounded-martingale inequality shows that empirical edge exposures divided by t differ from the average conditional exposures by at most $t^{-1/8}$, except with probability $A'' e^{-c'' t^{3/4}}$. Uniform continuity and the current mean-estimation event then give, for every fixed $\epsilon > 0$ and all large t, $\Pr\{Z_{a^*}(t) < t(\Gamma_M^{\mathrm{mat}} - \epsilon)\} \le A_\epsilon e^{-c_\epsilon t^{1/8}}$, where a^* is the stable answer or the selected optimal witness. On this event's complement the answer is empirically valid as well.

For any fixed $\eta > 0$, choose ϵ sufficiently small. Since $\lambda_\delta \to 1$ and the additional terms in β_{loc} are $O(\log P + \log\log(1/\delta) + \log(t+1))$, for all sufficiently small δ and every $t \ge (1+\eta) T_M^{\mathrm{mat}} \log(1/\delta)$, one has $\beta_{\mathrm{loc}}(t,\delta) < t(\Gamma_M^{\mathrm{mat}} - \epsilon)$. The ratio $\beta_{\mathrm{loc}}(t,\delta)/t$ is eventually decreasing, so the assertion holds for the entire tail. Summing the displayed summable bound on $\Pr\{\tau_\delta > t\}$ proves the expected-time upper limit $(1+\eta) T_M^{\mathrm{mat}}$. Let $\eta \downarrow 0$ and combine with the lower bound. The finite expectation also implies almost-sure termination. The argument fixes the dimension before taking $\delta \downarrow 0$; it is not a dimension-uniform finite-time leading-constant guarantee. □

For unrestricted scalar queries, a four-count version of the local union bound uses $6\log(t+1)$ in place of $5\log(t+1)$. The scalar construction below instead uses the Gaussian mixture boundary.

D.3 Scalar tracking

For scalar queries, use β_{sc} and forced exploration $t^{-1/2}\mathbf{1}/d$. Select the lexicographically first empirical blocker within ξ_t of the best estimated scalar rate. When no empirical blocker exists, maximize the stable objective regularized by $-\xi_t \|w\|_2^2$. The regularization makes the oracle single valued without changing its limiting value.

Theorem D.1 (Optimal scalar tracking)**.** *Algorithm 4 is δ-correct. For every strict audit instance,*

$$\limsup_{\delta \downarrow 0} \frac{\mathbb{E}_\theta[\tau_\delta]}{\log(1/\delta)} \le \begin{cases} T_M^{\mathrm{st}}, & M \textit{ stable}, \\ T_M^{\mathrm{un}}, & M \textit{ unstable}. \end{cases}$$

Together with (2)*, the algorithm is asymptotically instance optimal.*

Algorithm 4 Stable-audit Track-and-Stop

Require: proposed matching M, confidence δ

sample every coordinate once
for $t = d, d+1, \dots$ **do**
  compute empirical gains and $\widehat{\texttt{block}}_t$
  if $\widehat{\texttt{block}}_t = \varnothing$ **then**
    set the active answer to `stable` and compute the ξ_t-regularized stable allocation w_t
  else
    select the stabilized blocker p_t described above and compute w_t from Theorem 3.1
  end if
  if $Z_{\mathrm{st}}(t) \geq \beta_{\mathrm{sc}}(t, \delta)$ and $\widehat{\texttt{block}}_t = \varnothing$ **then**
    return `stable`
  else if $\widehat{\texttt{block}}_t \neq \varnothing$ and $Z_{p_t}(t) \geq \beta_{\mathrm{sc}}(t, \delta)$ **then**
    return p_t
  end if
  C-track the target $(1 - t^{-1/2})w_t + t^{-1/2}\mathbf{1}/d$ to choose the next coordinate
end for

Lemma D.2 (Gaussian mixture boundary)**.** *Assume every coordinate has been sampled at least once and fix $\eta > 0$. With probability at least $1 - \delta$, simultaneously for all t,*

$$L_t(\theta) < (1 + \eta)\left[\log(1/\delta) + \frac{d}{2}\log\left(1 + \frac{t}{d\eta}\right)\right].$$

Proof. Let $S_e(t)$ be the sum of centered observations from coordinate e. For every fixed vector $\lambda \in \mathbb{R}^d$, $\exp\{\sum_e \lambda_e S_e(t) - \frac{1}{2}\sum_e \lambda_e^2 N_e(t)\}$ is a nonnegative martingale under adaptive sampling. Integrate it with respect to independent Gaussian priors $\lambda_e \sim \mathcal{N}(0, 1/\eta)$. Fubini's theorem gives the nonnegative martingale

$$\mathcal{M}_t = \prod_e \sqrt{\frac{\eta}{N_e(t) + \eta}} \exp\left\{\frac{S_e(t)^2}{2(N_e(t) + \eta)}\right\},$$

with expectation one. Since $N_e(t) \geq 1$, $N_e/(N_e + \eta) \geq 1/(1 + \eta)$, and therefore

$$\log \mathcal{M}_t \geq \frac{L_t(\theta)}{1 + \eta} - \frac{1}{2}\sum_e \log \frac{N_e(t) + \eta}{\eta}.$$

Concavity of the logarithm and $\sum_e N_e(t) = t$ imply $\sum_e \log[(N_e + \eta)/\eta] \leq d \log[1 + t/(d\eta)]$. Ville's inequality gives $\mathbb{P}(\sup_t \mathcal{M}_t \geq 1/\delta) \leq \delta$. On the complementary event, rearranging the last two displays gives the result. □

For tracking, treat the initial d observations as deterministic unit target vectors. Restart the discrepancy accounting after this fixed initialization; its additive $O(d)$ offset does not affect any convergence or tail statement below.

Lemma D.3 (Tracking discrepancy)**.** *Let $\widetilde{w}_t \in \Delta_d$ be any predictable sequence and choose at round $t + 1$ a coordinate maximizing the deficit $\sum_{s \leq t+1} \widetilde{w}_{s,e} - N_e(t)$. Then, for every t,*

$$\max_e \left| N_e(t) - \sum_{s \leq t} \widetilde{w}_{s,e} \right| \leq d - 1.$$

If $\widetilde{w}_t = (1 - t^{-1/2})w_t + t^{-1/2}\mathbf{1}/d$, then $N_e(t) \geq (2\sqrt{t} - 2)/d - (d - 1)$ for every coordinate.

Proof. Put $P_e(k) = \sum_{s \leq k} \widetilde{w}_{s,e}$. We first prove by induction that $\max_e[N_e(k) - P_e(k)] \leq 1$. The claim is immediate at $k = 0$. If coordinate a is selected at round $k + 1$, then for $e \neq a$, $N_e(k+1) - P_e(k+1) = N_e(k) - P_e(k) - \widetilde{w}_{k+1,e} \leq 1$. For the selected coordinate,

$$N_a(k+1) - P_a(k+1) = 1 - [P_a(k+1) - N_a(k)] \leq 1,$$

because the selected bracket is the maximum of numbers whose sum is one and is therefore nonnegative.

It remains to control the opposite deviation. Define $r_k = \sum_e (P_e(k) - N_e(k))_+$. We prove $r_k \le d-1$ by induction. The claim holds at zero. Let a again be the selected coordinate. If $P_a(k+1) \ge N_a(k+1)$, updating the positive parts gives $r_{k+1} \le r_k + \sum_e \widetilde{w}_{k+1,e} - 1 = r_k$. Otherwise, the selected post-allocation deficit is negative. Then $P_a(k+1) - N_a(k) < 1$. Since a maximizes $P_e(k+1) - N_e(k)$, every unselected post-allocation deficit is at most this maximum, and hence

$$\begin{aligned} r_{k+1} &= \sum_{e \ne a} (P_e(k+1) - N_e(k+1))_+ \\ &\le \sum_{e \ne a} \max_j (P_j(k+1) - N_j(k)) < d-1. \end{aligned}$$

Thus $r_k \le d-1$ for all k. Since the coordinate errors sum to zero,

$$\max_e |N_e(k) - P_e(k)| \le \max\left\{ \max_e [N_e(k) - P_e(k)], r_k \right\} \le d-1$$

for $d \ge 2$; the case $d = 1$ is exact. Finally, $P_e(t) \ge d^{-1} \sum_{s \le t} s^{-1/2} \ge (2\sqrt{t} - 2)/d$, and subtracting the discrepancy proves the forced-exploration bound. □

Lemma D.4 (Forced-exploration concentration)**.** *For every fixed $r > 0$, there are constants $C_r, c_r > 0$ such that*

$$\mathbb{P}_\theta \left(\sup_{s \ge t} \|\widehat{\theta}(s) - \theta\|_\infty > r \right) \le C_r \exp(-c_r \sqrt{t}).$$

Moreover, almost surely $\|\widehat{\theta}(t) - \theta\|_\infty = o(t^{-1/8})$.

Proof. For coordinate e, expose in advance an infinite independent sequence $Y_{e,1}, Y_{e,2}, \dots$ of $\mathcal{N}(\theta_e, 1)$ observations. Whenever the algorithm samples e, it reveals the next unused element of this sequence. By Lemma D.3, every time $s \ge t$ uses at least $m_t = c\sqrt{t}$ samples of every coordinate for all sufficiently large t. If the empirical mean of coordinate e deviates by more than r at some such time, then for some integer $m \ge m_t$, the average of $Y_{e,1}, \dots, Y_{e,m}$ deviates by more than r. A union bound over m and the Gaussian tail give

$$\mathbb{P}\left(\exists m \ge m_t : \left| m^{-1} \sum_{k \le m} (Y_{e,k} - \theta_e) \right| > r \right) \le 2 \sum_{m \ge m_t} e^{-mr^2/2} \le C_r e^{-c_r \sqrt{t}}.$$

A union bound over the d coordinates proves the first claim.

For the second claim, apply the same argument at integer time t with threshold $t^{-1/8}/K$ for any fixed $K > 0$. The resulting probability is at most $C \exp(-ct^{1/4})$, which is summable. Borel–Cantelli yields $t^{1/8} \|\widehat{\theta}(t) - \theta\|_\infty \le 1/K$ eventually, for every positive integer K. Taking the countable intersection gives the asserted little-o rate. □

Lemma D.5 (Tracked information rate)**.** *Let $Z(t)$ denote the GLR statistic of the active answer. Almost surely,*

$$\frac{Z(t)}{t} \longrightarrow \begin{cases} 1/T_M^{\mathrm{st}}, & M \text{ stable}, \\ 1/T_M^{\mathrm{un}}, & M \text{ unstable}. \end{cases}$$

For every $\epsilon > 0$, there are constants $C_\epsilon, c_\epsilon, t_\epsilon > 0$ such that

$$\mathbb{P}_\theta \left(\frac{Z(t)}{t} < (1-\epsilon)/T \text{ or the active answer has not stabilized} \right) \le C_\epsilon \exp(-c_\epsilon t^{1/8})$$

for $t \ge t_\epsilon$, where T is the corresponding characteristic time.

Proof. Write $W^{\mathrm{st}}(\theta)$ for the compact convex maximizer set of the stable objective. We establish the rate claim through answer selection, oracle consistency, and tracking.

Strictness gives a positive minimum absolute gain among the finite set of current-versus-candidate comparisons. By Lemma D.4, after a random finite time the empirical blocker set equals the true blocker set; the tail of this time is at most $C\exp(-c\sqrt{t})$.

On an unstable instance, every true blocker has gains bounded away from zero in a neighborhood of θ, and the rate $[8(a^{-2}+b^{-2})]^{-1}$ is locally Lipschitz. The empirical rate error is therefore $o(t^{-1/8})$ almost surely. Suboptimal blockers are separated from the optimum by a fixed positive rate gap, so they eventually fail the $\xi_t = t^{-1/8}$ tolerance. Every rate-maximizing blocker eventually lies within that tolerance, and the lexicographic rule then selects the same first maximizer forever. The deviation estimate in the proof of Lemma D.4, used with a constant multiple of $t^{-1/8}$ and union-bounded over later times, bounds the stabilization tail by $C\exp(-ct^{1/4})$. On a stable instance the empirical blocker set is eventually empty, so the active answer is eventually `stable` with the stronger $\exp(-c\sqrt{t})$ tail.

On a stable strict instance, the empirical gains eventually have the same sign pattern as the true gains. The function $f(\vartheta, w) = \min_p I_p(w;\vartheta)$ is jointly continuous on a neighborhood of θ times the compact simplex. The regularized objective $f(\widehat{\theta}(t), w) - \xi_t\|w\|_2^2$ is strictly concave and has a unique maximizer. If a subsequence of these maximizers stayed a positive distance from $W^{\mathrm{st}}(\theta)$, compactness would give a convergent further subsequence. Uniform convergence of the regularized objectives to $f(\theta,\cdot)$ would make its limit a maximizer of $f(\theta,\cdot)$, a contradiction.

For a fixed blocker, the allocation in Theorem 3.1 is a continuous function of its two positive gains and is unique. Answer stabilization and converging gains therefore imply convergence of the unstable oracle allocation.

By Lemma D.3, the difference between $N(t)/t$ and the average of the mixed target weights is $O(d/t)$. The average forced-exploration mass is $t^{-1}\sum_{s\le t} s^{-1/2} = O(t^{-1/2})$. On a stable instance, the oracle argument above shows that the distance of w_s from the compact convex set $W^{\mathrm{st}}(\theta)$ tends to zero. Choose $v_s \in W^{\mathrm{st}}(\theta)$ closest to w_s. The average of the v_s lies in $W^{\mathrm{st}}(\theta)$ by convexity, while the average distance between v_s and w_s tends to zero by Cesaro convergence. Thus the empirical stable proportions approach the optimal set. On an unstable instance, the target allocations themselves converge to one vector, so their Cesaro averages and the tracked counts converge to that vector.

Almost-sure convergence follows from the preceding answer and allocation convergence arguments and continuity of the projection formulas in empirical gains and normalized positive counts. On a stable instance, every accumulation point of $N(t)/t$ belongs to $W^{\mathrm{st}}(\theta)$, where $\min_p I_p$ equals $1/T_M^{\mathrm{st}}$. On an unstable instance, Theorem 3.1 gives the limit at the selected optimal blocker.

For the tail statement, fix a target accuracy $\rho > 0$. On a stable instance, compactness and continuity give an objective gap: outside the ρ-neighborhood of $W^{\mathrm{st}}(\theta)$, the true stable objective is smaller than $1/T_M^{\mathrm{st}}$ by some $a_\rho > 0$. Uniform convergence of the empirical regularized objectives therefore puts every late stable oracle allocation inside that neighborhood; this argument does not require a single-valued or continuous unregularized optimizer. On an unstable instance, continuity of the explicit blocker allocation and the strict rate gap to suboptimal blockers give the analogous conclusion after answer stabilization. Discarding the first $\lfloor\sqrt{t}\rfloor$ targets changes the average allocation by $O(t^{-1/2})$. Thus, if every empirical parameter after time $\sqrt{t}$ has the sign and $s^{-1/8}$ accuracy required in the answer-selection argument above, C-tracking and continuity of the projection costs imply $Z(t)/t \ge (1-\epsilon)/T$ for all sufficiently large t. Lemma D.4 and a union bound over $s \ge \sqrt{t}$ bound the complementary event by $C_\epsilon\exp(-c_\epsilon t^{1/8})$.

□

Proof of Theorem D.1. Apply Lemma D.2 with $\eta = \eta_\delta$. If an answer is incorrect, the true parameter belongs to its alternative set, so $Z_a(t) \le L_t(\theta)$. An incorrect stop can therefore occur only on the probability-δ complement of the mixture event. Thus the algorithm is δ-correct.

Let T be the appropriate characteristic time and fix $\epsilon > 0$. Set $\mathcal{L}_\delta = \log(1/\delta)$ and $t_\delta = (1+\epsilon)T\mathcal{L}_\delta$. Choose $\rho > 0$ so small that $(1+\epsilon)(1-\rho) > 1 + \epsilon/2$. Since $\eta_\delta = 1/\log(e/\delta)$,

$$\sup_{t \in [t_\delta, 2t_\delta]} \frac{\beta_{\mathrm{sc}}(t,\delta)}{\mathcal{L}_\delta} = 1 + o(1).$$

Hence, for all sufficiently small δ, $t_\delta(1-\rho)/T > \beta_{\mathrm{sc}}(t_\delta, \delta)$. Moreover $\partial_t \beta_{\mathrm{sc}}(t,\delta) = (1+\eta_\delta)d/[2(d\eta_\delta + t)]$, which is smaller than $(1-\rho)/T$ for every $t \geq t_\delta$ once δ is small. Thus a stabilized path with $Z(t)/t \geq (1-\rho)/T$ stops at that checkpoint for every $t \geq t_\delta$. Therefore Lemma D.5 gives

$$\mathbb{P}_\theta(\tau_\delta > t) \leq C \exp(-ct^{1/8}), \qquad t \geq t_\delta,$$

with instance-dependent constants independent of sufficiently small δ. Integrating the tail,

$$\mathbb{E}_\theta \tau_\delta \leq t_\delta + \sum_{t \geq t_\delta} C e^{-ct^{1/8}} = t_\delta + o(\mathcal{L}_\delta).$$

Divide by $\mathcal{L}_\delta$, let $\delta \downarrow 0$, and then let $\epsilon \downarrow 0$ to obtain the upper bound. The matching lower bound is (2). □

E Proof and implementation of the information certificate

E.1 Statistical guarantee

We prove Theorem 5.1 without assuming that empirical endpoint signs are correct. The information formula in the matching-feedback model is a sum, over the two endpoints, of a nonnegative harmonic exposure coefficient times one half of the squared protection deficit. Thus for any nonnegative deficit arrays $d^- \leq \mathsf{d} \leq d^+$, $J_p(x; d^-) \leq J_p(x; \mathsf{d}) \leq J_p(x; d^+)$ for every x. The artificial arrays $d^\pm$ need not themselves arise from one realizable utility profile; monotonicity and the concave homogeneous information functions remain valid for these arrays.

For each of the $d = 2n^2$ utility streams, pre-generate an infinite independent sequence of unit-variance Gaussian observations. At deterministic prefix k, the probability that its empirical mean differs by more than $r(k) = \sqrt{2\log(\pi^2 dk^2/(3\eta))/k}$ is at most $6\eta/(\pi^2 dk^2)$. Summation over all streams and all positive integers k is at most η. Adaptive matching selection only chooses which prefixes are inspected; it does not require conditioning the Gaussian distribution on the chosen prefix. At count zero use an unbounded interval, or first deploy the cyclic initialization.

A pair gain is the candidate utility minus its current partner utility. Subtract the corresponding coordinate intervals, using their actual counts, to obtain $[\ell_p^s, u_p^s]$. On the event just proved, $[-u_p^s]_+ \leq [-g_p^s]_+ \leq [-\ell_p^s]_+$. If $L^- > 0$, every pair has a strictly positive certified deficit at an endpoint. The matching is therefore stable even if other endpoint intervals straddle zero.

For any chosen x, its true information rate is at least L^-. The true optimal rate is at most the optimum of the artificial upper-deficit game. For any chosen pair mixture λ, concavity and Euler's identity give $\min_p J_p(y; d^+) \leq \langle G^+, y \rangle$. Its maximum over the doubly stochastic polytope equals the maximum-weight assignment U^+. Consequently $L^- \leq \Gamma_M^{\mathrm{mat}}(\theta) \leq U^+$ and the actual efficiency is at least L^-/U^+. This implication is simultaneous for every feasible x, λ on the same event. Trying several solvers, comparing their outputs, or inspecting this ratio at data-dependent times incurs no additional statistical multiplicity cost. A deterministic relative optimization tolerance, however, need not vanish and is not by itself a proof of the tracking theorem's asymptotic leading constant.

For implementation, a proposed assignment dual with $\alpha_i + \nu_j \geq G_{ij}^+$ gives a further valid upper bound $\sum_i \alpha_i + \sum_j \nu_j$; exact feasibility or outward error control is required for a machine-rigorous numerical certificate. The supplied calculations check floating-point feasibility to disclosed tolerances, rather than claiming interval-arithmetic verification.

The script `run_statistical_design_certificates.py` uses thirty paths with eleven geometrically spaced common per-edge counts from 64 through 65,536. Gaussian sufficient-statistic increments have the exact joint prefix law. At each look it compares uniform exposure with a numerical empirical design, retaining the better certified ratio; 520 design comparisons are evaluated

at the admissible looks. All thirty paths eventually pass the one-half efficiency threshold. The reference true-mean design is used only for checking the inequalities, not for selecting the reported certificate. All looks and confidence-event indicators are in `data/statistical_design/statistical_design_certificates.json`.

E.2 Gradients, assignment bounds, and numerical checks

Relabel the proposed matching as the identity and write the stable pair rate as $J_{ij}(x) = a_{ij}h(x_{ij}, x_{ii}) + b_{ij}h(x_{ij}, x_{jj})$, where $h(u, v) = uv/(u + v)$ and a_{ij}, b_{ij} are one half the squared negative-gain parts on the two sides. On positive exposures, $\partial_u h = v^2/(u + v)^2$ and $\partial_v h = u^2/(u + v)^2$. Thus each pair gradient is supported on only three physical edges. The functions are differentiable, concave, and degree-one homogeneous. Euler's identity gives $J_p(x) = \langle \nabla J_p(x), x\rangle$.

The tangent-assignment argument above remains valid for any pair mixture, including a prematurely stopped dual optimizer. Positivity avoids ambiguous boundary derivatives and can be ensured by a vanishing mixture with uniform exposure.

For fixed x, the smallest such tangent bound can be found by minimizing $\sum_i \alpha_i + \sum_j \nu_j$ over $\lambda \geq 0$, $\sum_p \lambda_p = 1$, and $\alpha_i + \nu_j \geq \sum_p \lambda_p(\nabla J_p(x))_{ij}$. The potentials are unrestricted real numbers. This is a linear program (LP) whose sparse coefficients are determined by the three-edge pair gradients. Once a feasible distribution λ is obtained, recomputing one maximum-weight assignment from its actual gradient avoids trusting the reported LP objective. If homogeneity is evaluated numerically, one may retain the constant $\sum_p \lambda_p J_p(x) - \langle G, x\rangle$ instead of setting it to zero. A true mathematical or interval certificate also requires feasible exposures, nonnegative normalized weights, and conservative arithmetic for rates and derivatives. Solver termination flags alone establish none of these facts.

On a truly stable instance, let $L_t^{\mathrm{cert}}, U_t^{\mathrm{cert}}$ be the certificate bounds computed with empirical deficits. Suppose a tracking scheme returns positive designs with $U_t^{\mathrm{cert}} \leq (1 + \varepsilon_t)L_t^{\mathrm{cert}}$ for deterministic $\varepsilon_t \to 0$. Eventually the empirical instance is strictly stable. The certified lower rate is then at least the empirical optimum divided by $1 + \varepsilon_t$. The continuity and compactness argument used for the original tracking result therefore places these designs asymptotically in the true optimal design set; forced exploration and the same expected-tail proof remain necessary. This statement is an oracle replacement criterion, not a promise that any fixed-tolerance numerical solver satisfies it. Named-blocker swap designs retain their separate explicit optimum.

The numerical check uses sequential least-squares programming (SLSQP) only to propose an exposure matrix. A Birkhoff decomposition followed by normalized mixture reconstruction yields the deployed candidate, with an additional uniform mass of 10^{-7}. A separate LP chooses blocker weights, and a separate assignment solve evaluates the upper bound. Forty candidate instances comprise four seeds for each combination of five market sizes and two protection families. Each candidate has a relative gap below 0.0005 in floating-point arithmetic; exposure residuals are below $5 \cdot 10^{-16}$. Fifty independently sampled feasible mixtures per instance satisfy the upper bound. An exact rational-arithmetic two-market check verifies 171 rational tangents against the exact homogeneous optimum. These tests do not prove the precision of all floating-point certificates, and their timings are not substituted for end-to-end sequential stopping time.

F Protection structure and star analysis

F.1 The blocker-load dual

Proof of Theorem 3.3. Fix a star and suppress its index. Its objective is $F(c, x) = \sum_l b_l h(c, x_l)$. For a nonzero load the following calculation solves its unit-budget maximum.

This function is continuous and concave on the unit simplex, so a maximizer exists. Coordinates with $b_l = 0$ may be assigned zero mass. For a positive coefficient, the one-sided derivative with respect to x_l at zero is b_l; hence a leaf is active exactly when this derivative exceeds the simplex multiplier.

Let the multiplier be $\lambda = s^2$. For an active leaf, stationarity gives $b_l c^2/(c + x_l)^2 = s^2$, so $x_l = c(\sqrt{b_l}/s - 1)$ and activity is equivalent to $\sqrt{b_l} > s$. Stationarity in the center coordinate gives $\sum_l b_l x_l^2/(c + x_l)^2 = s^2$. Substituting the active-leaf formula yields $s^2 = \sum_l (\sqrt{b_l} - s)_+^2$. The left

side is strictly increasing in s, whereas the right side is continuous and strictly decreasing until it reaches zero. At $s = 0$ the right side is positive, and at $s = \max_l \sqrt{b_l}$ it is zero, so the solution is unique. Normalizing the masses gives the displayed c and x_l. These Karush–Kuhn–Tucker (KKT) conditions are sufficient by concavity.

The objective is positively homogeneous of degree one in (c, x). Euler's identity and stationarity on the active coordinates therefore give $F(c, x) = \lambda(c + \sum_l x_l) = s^2$. Finally, $\Lambda(b)$ is the pointwise maximum, over feasible masses, of functions linear in b; it is consequently convex and coordinatewise nondecreasing. Positive homogeneity in b is immediate from the definition.

Within star r, collecting equal leaves gives $F_r(c_r, x_r; q) = \sum_l b_{rl} h(c_r, x_{rl})$, where $b_{rl} = \rho_{rl}(q)\Delta_{rl}^2/2$. The utility coordinates of different stars are disjoint. By positive homogeneity of h, rescaling a unit-mass optimizer by m_r multiplies its value by m_r, and no allocation of mass m_r can do better after normalization. Thus the optimal contribution is $m_r \Lambda_r(q)$.

The stars have disjoint utility coordinates. For fixed q, distributing total masses m_r therefore gives inner optimum $\max_{m \in \Delta(\mathcal{R})} \sum_r m_r \Lambda_r(q) = \max_r \Lambda_r(q)$. Minimize over q and apply Theorem 3.1. A doubly protected pair contributes its mass to both relevant stars. Convexity follows because each local optimum is a maximum of linear functions of the loads.

□

Theorem F.1 (Pressure optimality certificate)**.** *A blocker law q^* minimizes the blocker-load dual if and only if it belongs to a blocker-pressure equilibrium with $\gamma = \max_r \Lambda_r(q^*)$. Scaling each active star's unique local optimizer by μ_r produces a primal allocation w^* satisfying $I_p(w^*) = G_p$. Hence (w^*, q^*) is a saddle point, and the equal-pressure conditions are an exact, checkable optimality certificate.*

Proof of Theorem F.1. Write the blocker-load problem in epigraph form as

$$\min_{q,t} \quad t \qquad \text{subject to} \qquad q \in \Delta(\mathcal{P}_M), \quad \Lambda_r(q) \le t \quad (r \in \mathcal{R}).$$

The functions Λ_r are convex by Theorem 3.3. The problem has a relative-interior Slater point: choose q with full support and then choose t strictly larger than every local load. Hence its Karush–Kuhn–Tucker conditions are necessary and sufficient.

Whenever $\Lambda_r(q^*) > 0$, the loaded-star optimizer is unique. The scalar s in Theorem 3.3 is unique, its active set is determined by $\sqrt{b_{rl}} > s$, and normalization uniquely fixes the center and every leaf, with inactive leaves equal to zero. Danskin's theorem therefore gives

$$\frac{\partial \Lambda_r(q^*)}{\partial q_p} = \sum_{l : p(r,l) = p} \frac{\Delta_{rl}^2}{2} h(c_r(q^*), x_{rl}(q^*)) = g_{rp}(q^*)$$

for every active epigraph constraint, where $p(r, l)$ is the candidate using leaf l of star r. Stars with zero multiplier need not be differentiated.

Let $\mu_r \ge 0$ be the multiplier of $\Lambda_r(q) \le t$, let ν be the multiplier of $\sum_p q_p = 1$, and let $\xi_p \ge 0$ be the multiplier of $q_p \ge 0$. Stationarity and complementary slackness give

$$\sum_r \mu_r = 1, \qquad \mu_r(\Lambda_r(q^*) - t^*) = 0,$$
$$\sum_r \mu_r g_{rp}(q^*) + \nu - \xi_p = 0, \qquad \xi_p q_p^* = 0.$$

Consequently, $G_p = -\nu$ on $\operatorname{supp}(q^*)$ and $G_p \ge -\nu$ off the support. Moreover,

$$\sum_p q_p^* G_p = \sum_r \mu_r \sum_p q_p^* g_{rp}(q^*) = \sum_r \mu_r \Lambda_r(q^*) = t^*.$$

The left side equals $-\nu$ because q^* is supported on equality coordinates. Hence $-\nu = t^* = \gamma^*$, proving the displayed pressure conditions.

Conversely, suppose q^* and μ satisfy those conditions. Scale the unit local optimizer of star r by μ_r and concatenate the disjoint coordinates. Positive homogeneity of h gives $I_p(w^*) = \sum_r \mu_r g_{rp}(q^*) = G_p \geq \gamma^*$ for every p, so w^* attains primal value at least γ^*. On the other hand, for any allocation w,

$$\sum_p q_p^* I_p(w) \leq \max_r \Lambda_r(q^*) = \gamma^*$$

by Theorem 3.3. Thus no primal allocation can have minimum blocker information exceeding γ^*. Equality follows, and the support equalities show that q^* is a best response to w^*. This proves both optimality directions and the saddle statement. □

F.2 Disjoint components and the Gaussian star

Proof of Theorem 3.4. The blocker edges and utility coordinates of distinct protection components are disjoint. If component r receives total sampling mass m_r, positive homogeneity of every information function makes its best worst-edge rate m_r/T_r. The global primal therefore reduces to $\max_{m \in \Delta_m} \min_r m_r/T_r$. Its value is at most $1/\sum_r T_r$, because having every ratio larger would imply $\sum_r m_r > 1$. Equality holds at $m_r = T_r/\sum_s T_s$, proving additivity.

For the dual claim, let an arbitrary blocker law assign component r total probability Q_r, and condition it inside that component when $Q_r > 0$. By the definition of T_r, the best expected information available inside component r is at least Q_r/T_r after the adversary minimizes its conditional law. Hence every global dual law has value at least $\max_r Q_r/T_r \geq 1/\sum_s T_s$. Equality in the last inequality forces $Q_r = T_r/\sum_s T_s$ for every r; equality in the game also forces each conditional law to be locally least favourable. Conversely, combining any local least-favourable laws with these component probabilities gives expected information at most $\sum_r Q_r m_r/T_r = 1/\sum_s T_s$ against every global allocation with component masses m_r. Thus the stated component masses are necessary and sufficient on both sides of every saddle point. □

Isolated coordinates incident to no protecting comparison can be discarded. The components in Theorem 3.4 are the nonempty comparison components. For singly protected audits, assigning each pair to its unique rejecting endpoint makes these components disjoint stars; hence their times and optimal masses follow directly from the component theorem.

The following allocation is the Gaussian best-arm star calculation (Garivier & Kaufmann, 2016), expressed in audit notation.

Theorem F.2 (Gaussian star allocation)**.** *The unit-budget characteristic time of a protection star is*

$$H(\Delta_1, \ldots, \Delta_k) = \min_{z > \max_l a_l} \left\{ z + \sum_{l=1}^{k} \frac{a_l z}{z - a_l} \right\}.$$

The minimizer z^* is unique and satisfies $1 = \sum_l a_l^2/(z^* - a_l)^2$. Writing $H = H(\Delta_1, \ldots, \Delta_k)$, the unique optimal weights are $w_c = z^*/H$ and $w_l = a_l z^*/[H(z^* - a_l)]$. A least-favourable blocker law is $q_l \propto a_l/(z^* - a_l)^2$.

Proof. Put $F(z) = z + \sum_l a_l z/(z - a_l)$. We first prove the upper bound on every feasible information rate. Suppose center and leaf masses $(c, x_1, \ldots, x_k)$ have total at most one and every leaf receives information at least $\gamma > 0$. Since $I_l = h(c, x_l)/a_l$, the lth constraint is $cx_l/(c + x_l) \geq a_l\gamma$. It forces $c > a_l\gamma$ and, after rearrangement, $x_l \geq a_l\gamma c/(c - a_l\gamma)$. With $z = c/\gamma$, necessarily $z > \max_l a_l$, and summing the forced masses gives

$$1 \geq c + \sum_l x_l \geq \gamma \left\{ z + \sum_l \frac{a_l z}{z - a_l} \right\} = \gamma F(z).$$

Thus $\gamma \leq 1/\min_z F(z)$ for every allocation.

Conversely, let z^* minimize F and set $H = F(z^*)$, $c = z^*/H$, and $x_l = a_l z^*/[H(z^* - a_l)]$. These masses are positive and sum to one. Direct substitution gives $h(c, x_l) = a_l/H$ for every leaf, so every blocker has information $1/H$. Hence the upper bound is attained and the characteristic time is exactly H.

Using $a_l z/(z-a_l) = a_l + a_l^2/(z-a_l)$ gives

$$F'(z) = 1 - \sum_l \frac{a_l^2}{(z-a_l)^2}, \qquad F''(z) = 2\sum_l \frac{a_l^2}{(z-a_l)^3} > 0.$$

Hence F is strictly convex. It diverges as z decreases to $\max_l a_l$ and also as $z \to \infty$. Its derivative increases continuously from $-\infty$ to 1, so it has one zero, which is the unique minimizer.

The weights sum to one by the definition of H. Set $q_l = Ca_l/(z^* - a_l)^2$, with C normalizing the probabilities. The preceding direct substitution gives $I_l(w) = 1/H$ for every l, so q is a best response to w.

It remains to show that w maximizes $\sum_l q_l I_l(w)$ over the simplex. This objective is concave. Let $c = w_c$ and $x_l = w_l$. Since $I_l = h(c, x_l)/a_l$,

$$\frac{\partial I_l}{\partial x_l} = \frac{1}{a_l}\frac{c^2}{(c+x_l)^2}, \qquad \frac{\partial I_l}{\partial c} = \frac{1}{a_l}\frac{x_l^2}{(c+x_l)^2}.$$

At the displayed weights, $c/(c+x_l) = (z^* - a_l)/z^*$ and $x_l/(c+x_l) = a_l/z^*$. Therefore $q_l \partial I_l/\partial x_l$ is the same constant for every l. The center derivative is

$$\sum_l q_l \frac{a_l}{(z^*)^2}.$$

Writing $q_l = Ca_l/(z^* - a_l)^2$, this equals $C(z^*)^{-2}\sum_l a_l^2/(z^* - a_l)^2 = C/(z^*)^2$ by the optimality equation. The leaf derivatives equal the same constant $C/(z^*)^2$. Thus the KKT stationarity conditions on the simplex hold, and concavity makes them sufficient. Hence w is a best response to q. □

Equal deficits. For $a_l = a = 2/\Delta^2$, the stationarity equation gives $z^* = a(1+\sqrt{k})$. Thus $H = 2(1+\sqrt{k})^2/\Delta^2$, the center mass is $1/(1+\sqrt{k})$, each leaf mass is $1/[\sqrt{k}(1+\sqrt{k})]$, and the blocker law is uniform. The shared center receives $\sqrt{k}$ times the mass of one leaf.

F.3 Envelope and dominant leaves

Theorem F.3 (Effective degree and blocker concentration)**.** *Every protection star satisfies*

$$a_{\max}\max\{4, 1+k_{\text{eff}}\} \le H \le a_{\max}(1+\sqrt{k_{\text{eff}}})^2.$$

The upper bound is exact for equal deficits, while both bounds converge to $4a_{\max}$ when one inverse-squared deficit dominates all others. If leaf 1 has $a_1 = a_{\max}$, then its optimal blocker probability satisfies $q_1^* \ge a_1/A_\star$.

Proof. Let $F(z) = z + \sum_l a_l z/(z-a_l)$. Since every summand is at least a_l and $z > a_{\max}$, $F(z) > A_\star + a_{\max}$ for all feasible z. Also, deleting every leaf except one with coefficient $a_{\max}$ can only decrease the characteristic time. A one-leaf star has complexity $4a_{\max}$ by Theorem F.2. These two arguments give $H \ge \max\{A_\star + a_{\max}, 4a_{\max}\} = a_{\max}\max\{1+k_{\text{eff}}, 4\}$.

For the upper bound, $z - a_l \ge z - a_{\max}$ implies

$$F(z) \le z + \frac{A_\star z}{z - a_{\max}}.$$

The right side is minimized at $z = a_{\max} + \sqrt{A_\star a_{\max}}$ and its minimum is $(\sqrt{A_\star} + \sqrt{a_{\max}})^2 = a_{\max}(1+\sqrt{k_{\text{eff}}})^2$. If all $a_l = a_{\max}$, every denominator comparison used above is an equality and the scalar minimizer is the same, so the upper bound is exact. If $A_\star/a_{\max} \to 1$, the lower and upper bounds both converge to $4a_{\max}$.

Put $B = \sum_{l\ge 2} a_l$. Let z^* be the scalar minimizer. By Theorem F.2, the unnormalized dual weight of leaf l is $a_l/(z^* - a_l)^2$. Since $a_l \le a_1$ and $z^* - a_l \ge z^* - a_1$ for every $l \ge 2$,

$$\sum_{l=2}^{k} \frac{a_l}{(z^* - a_l)^2} \le \frac{B}{(z^* - a_1)^2}.$$

The first leaf has unnormalized weight $a_1/(z^* - a_1)^2$, so normalization yields $q_1^* \ge a_1/(a_1 + B)$.

□

Consequences for reuse. The envelope implies $4a_1 \le H \le 4a_1 + 2\sum_{l\ge 2} a_l$, using $\sqrt{1+x} \le 1 + x/2$. Hence a dominant leaf concentrates both the time and the blocker law. Separately, auditing leaves with separate center copies costs $4A_\star$, so $1 \le R_{\mathrm{reuse}} < 4$; the strict upper bound follows from $H > A_\star$. With equal deficits, $R_{\mathrm{reuse}} = 4k/(1+\sqrt{k})^2 \to 4$. The time envelope and the concentration of the blocker law describe the same dominant-leaf regime.

F.4 Sensitivity and curvature

Theorem F.4 (Differential geometry of the star cost)**.** *On the positive orthant $(0,\infty)^k$, $H(a)$ is smooth, symmetric, positively homogeneous, and convex. Its Hessian satisfies*

$$v^\top \nabla^2 H(a) v = 2(z^*)^2 \left[\sum_l \frac{v_l^2}{t_l^3} - \frac{(\sum_l a_l v_l / t_l^3)^2}{S} \right] \ge 0,$$

with equality exactly for radial directions $v \propto a$. Its first derivative is $\partial H/\partial a_l = [z^*/(z^* - a_l)]^2$. Its blocker leverage vector is exactly the least-favourable blocker law: $\ell_l = q_l^* = a_l (z^*)^2/[H(z^* - a_l)^2]$. In particular, $\sum_l \ell_l = 1$, and a smaller protecting gap has larger leverage.

Proof. Write $F(z,a) = z + \sum_l a_l z/(z - a_l)$. By Theorem F.2, the unique optimizer $z^*(a)$ is characterized by $F_z(z^*, a) = 0$, and

$$F_{zz}(z^*, a) = 2\sum_l \frac{a_l^2}{(z^* - a_l)^3} = 2S > 0.$$

The implicit-function theorem therefore makes z^*, and hence $H(a) = F(z^*(a), a)$, smooth on the positive orthant. Symmetry and positive homogeneity are immediate from the variational formula.

The required partial derivatives of F are

$$F_{a_l} = \frac{z^2}{(z - a_l)^2}, \qquad F_{a_l a_r} = \mathbf{1}\{l = r\} \frac{2z^2}{(z - a_l)^3},$$
$$F_{z a_l} = -\frac{2 a_l z}{(z - a_l)^3}, \qquad F_{zz} = 2\sum_j \frac{a_j^2}{(z - a_j)^3}.$$

Differentiating the envelope identity $H_{a_l} = F_{a_l}(z^*(a), a)$ and using $\partial z^*/\partial a_r = -F_{z a_r}/F_{zz}$ gives

$$H_{a_l a_r} = F_{a_l a_r} - \frac{F_{a_l z} F_{z a_r}}{F_{zz}},$$

which is the displayed Hessian formula.

For the quadratic form, substitute the Hessian entries and collect terms. The bracket is nonnegative by the Cauchy–Schwarz inequality applied to the vectors $(v_l/t_l^{3/2})_l$ and $(a_l/t_l^{3/2})_l$. Equality in Cauchy–Schwarz holds exactly when $v_l/t_l^{3/2} = c a_l/t_l^{3/2}$ for every l, equivalently $v = ca$. This proves convexity and the strictness assertion.

The envelope identity already gives $H_{a_l} = (z^*)^2/(z^* - a_l)^2$. Euler's identity consequently gives $\sum_l a_l \partial H/\partial a_l = H$. Hence

$$\ell_l = \frac{a_l}{H}\frac{\partial H}{\partial a_l} = \frac{a_l (z^*)^2}{H(z^* - a_l)^2}, \qquad \sum_l \ell_l = 1.$$

The normalization identity shows that this vector equals the law in Theorem F.2. Finally, for fixed z^* the map $a \mapsto a/(z^* - a)^2$ is strictly increasing on $(0, z^*)$. Thus $a_l \ge a_r$, equivalently $\Delta_l \le \Delta_r$, implies $\ell_l \ge \ell_r$.

□

Consequences for gap perturbations and majorization. Since $da_l/a_l = -2\,d\Delta_l/\Delta_l$, the leverage identity gives $dH/H = -2\sum_l q_l^* d\Delta_l/\Delta_l$. Symmetric convexity also implies $H(a) \geq H(b)$ whenever a majorizes b. At fixed $\sum_l a_l = A_\star$, equal coefficients minimize complexity at $A_\star(1+\sqrt{k})^2/k$, and a one-leaf vector maximizes its continuous extension at $4A_\star$. Here is the direct argument for the latter claims.

The formula in Theorem F.2 extends continuously to nonnegative a by deleting zero-coefficient leaves. Because H is symmetric and convex, Jensen's inequality applied to all coordinate permutations gives

$$H(A_\star \mathbf{1}/k) = H\left(\frac{1}{k!}\sum_\pi \pi a\right) \leq \frac{1}{k!}\sum_\pi H(\pi a) = H(a).$$

The equal-deficit formula gives $H(A_\star \mathbf{1}/k) = A_\star(1+\sqrt{k})^2/k$. If a is positive and not uniform, the averaged permutation directions are not all positive rescalings of one another; the strictness statement in Theorem F.4 makes the Jensen inequality strict.

Every a with total $A_\star$ is a convex combination of the k one-leaf vectors $A_\star e_l$, where e_l is the lth standard basis vector. Convexity and symmetry therefore yield $H(a) \leq \sum_l (a_l/A_\star) H(A_\star e_l) = 4A_\star$. The one-leaf formula shows equality at every vertex. Finally, if a majorizes b, the Hardy–Littlewood–Pólya characterization writes b as a convex combination of permutations of a. Convexity and symmetry then give $H(b) \leq H(a)$.

F.5 Value of retaining both endpoints

Property F.5 (Two-sided protection)**.** *The full two-sided audit rate obeys* $\Gamma_M^{\mathrm{or}} \leq \Gamma_M^{\mathrm{st}} \leq 2\Gamma_M^{\mathrm{or}}$. *Thus the protection redundancy index* $R_M = \Gamma_M^{\mathrm{st}}/\Gamma_M^{\mathrm{or}}$ *always lies in* $[1,2]$.

Proof of Property F.5. For every candidate pair p, let C_p be its set of negatively signed endpoint comparisons. An orientation σ chooses one member of C_p and retains only its information contribution $I_{p,\sigma}(w)$. Since $I_p(w) = \sum_{c\in C_p} I_{p,c}(w) \geq I_{p,\sigma}(w)$, every orientation satisfies $\Gamma_M^{\mathrm{st}} \geq \Gamma_{M,\sigma}$, and maximizing over orientations gives the first inequality.

Conversely, fix an allocation w and orient each p toward a largest endpoint contribution. Since $|C_p| \leq 2$, $I_p(w) \leq 2I_{p,\sigma_w}(w)$ for every p. Therefore

$$\min_p I_p(w) \leq 2\min_p I_{p,\sigma_w}(w) \leq 2\max_\sigma \max_{w'\in\Delta_d} \min_p I_{p,\sigma}(w') = 2\Gamma_M^{\mathrm{or}}.$$

Maximizing the left side over w proves the upper bound. □

G Homogeneous examples and normalized dimensions

G.1 The examples in Table 1

Take the identity matching. For every $i \in A$, set $u_{ii} = \Delta$ and $u_{ij} = 0$ for $j \neq i$. For every $j \in B$, set $v_{jj} = 0$ and $v_{ji} = \Delta$ for $i \neq j$. Ties among off-match alternatives may be broken exogenously; all current-versus-candidate gains are nonzero. Every candidate pair is preferred by its B endpoint and rejected by its A endpoint with deficit Δ, so the identity is stable and the protection hypergraph consists of n equal-gap stars with $k = n-1$ leaves. Theorems 3.4 and F.2 gives $T_M^{\mathrm{st}} = 2n(1+\sqrt{n-1})^2/\Delta^2$.

For the unstable instance, choose one pair (i,j) and replace $u_{ij} = 0$ by $u_{ij} = 2\Delta$, leaving all other coordinates unchanged. That pair is the unique blocker and its two gains are Δ. By Theorem 3.1, its characteristic time is $16/\Delta^2$. Their scalar-query ratio is $n(1+\sqrt{n-1})^2/8$.

The matching-round calculation is in Appendix B. Its row-star optimizer is also column feasible, so parallel exposure removes exactly the factor n from the stable scalar time.

G.2 Symmetric double protection

If every off-diagonal pair is rejected by both endpoints with common deficit Δ, the stable scalar time remains $2n(1+\sqrt{n-1})^2/\Delta^2$. The following symmetry calculation also gives its allocation.

The stable audit objective is concave in the allocation and invariant under simultaneous relabeling of agents, under relabeling the two market sides, and under permutations of off-diagonal coordinates within the induced orbits. Averaging any optimal allocation over this finite symmetry group cannot decrease the minimum blocker information. Thus an optimum exists with a common weight c on each of the $2n$ current-partner coordinates and a common weight x on each of the $2n(n-1)$ off-diagonal coordinates.

The simplex constraint is $2nc + 2n(n-1)x = 1$. Every candidate pair has two equal protecting contributions, so its information is $\Delta^2 h(c, x)$. Put $k = n - 1$ and $B = 1/(2n)$. Maximizing $h(c, x)$ subject to $c + kx = B$ is the scaled equal-gap star problem. Its unique solution is $c = B/(1 + \sqrt{k})$ and $x = B/[\sqrt{k}(1 + \sqrt{k})]$, and its value is $B/(1 + \sqrt{k})^2$. Hence $\Gamma_M^{\mathrm{st}} = \Delta^2/[2n(1 + \sqrt{n-1})^2]$. Inversion gives the characteristic time and the displayed coordinate weights.

G.3 Margin-normalized interpretation

The rate and margin results immediately imply $2 \le D_p^{\mathrm{un}} \le 4$ and

$$\frac{n(1+\sqrt{n-1})^2}{2} \le D_n^{\mathrm{st}} \le 2n^2.$$

Normalization removes the cardinal utility scale: a witness has constant dimension, while a universal certificate can have dimension of order n^2.

For a true blocker with gains $a, b > 0$, let $m = \min\{a, b\}$. By Theorem 3.1,

$$D_p^{\mathrm{un}} = \frac{m^2}{4}\, 8(a^{-2} + b^{-2}) = 2m^2(a^{-2} + b^{-2}).$$

One of a, b equals m and the other is at least m, so the last quantity lies in $[2, 4]$.

For a stable audit, Theorem 3.2 gives $T_M^{\mathrm{st}} \le d/\kappa^2_{\Psi_M^{\mathrm{aud}}}$ with $d = 2n^2$, proving $D_n^{\mathrm{st}} \le 2n^2$. In the stable construction of Table 1, every candidate pair has audit radius $\Delta/2$. Hence

$$D_M^{\mathrm{st}} = \frac{\Delta^2}{4}\frac{2n(1+\sqrt{n-1})^2}{\Delta^2} = \frac{n(1+\sqrt{n-1})^2}{2},$$

which proves the lower bound.

H Target geometry and separation examples

Proposition H.1 (Target-changing faces and irrelevant gaps)**.** *For every stable-market target, $\mathfrak{D}_\Psi$ is a finite union of faces of the tie arrangement and $\mathfrak{D}_\Psi \subseteq \mathfrak{H}$. Equality holds for Ψ_{prof}. For M^A, G, and $\mathcal{S}$, the inclusion is strict for $n \ge 3$. For every $n \ge 3$, every $\varepsilon > 0$, and every $g > 0$, there is a strict market whose minimum preference gap is at most ε but for which $\kappa_{M^A}(\theta)$, $\kappa_G(\theta)$, and $\kappa_{\mathcal{S}}(\theta)$ are all at least $g/2$.*

Proof. The arrangement induces a finite polyhedral complex whose full-dimensional cells are the closures of the connected components of $\mathbb{R}^d \setminus \mathfrak{H}$. Let F be any face of this complex. The set of full-dimensional cells containing F is finite. If two such cells have different Ψ-values, then every relative-interior point of F has arbitrarily close strict points with different values, and the same is true for every subface of F. If all incident full-dimensional cells have the same value, a sufficiently small ball around any relative-interior point of F meets only cells with that value, so that point is not in the boundary under consideration. Taking the union of the first class of faces gives the claimed set. No point outside $\mathfrak{H}$ belongs to it because every such point lies in the interior of one component.

Apply this arrangement argument to within-agent ties. Each chamber fixes the preference profile, so every stable-market target is constant on it. The full profile changes across every tie face and has discriminant $\mathfrak{H}$. The following construction proves both strict inclusion for the other three targets and the quantitative irrelevant-gap claim.

Choose $\xi \in (0, \min\{\varepsilon, g/4\})$. For every $i \in A$, set $u_{ii} = 2g$ and choose distinct values for $\{u_{ij} : j \ne i\}$ inside $[0, g]$ with two consecutive values at distance ξ. Make the analogous construction

for every $j \in B$, with $v_{jj} = 2g$ and all tail values in $[0, g]$. The minimum preference gap is at most $\xi \leq \varepsilon$.

Every identity pair is mutual top. Therefore any stable matching must match i to i for every i, since otherwise (i, i) is a blocking pair. The identity is the unique stable matching. Now perturb every utility coordinate by less than $g/2$ in absolute value. Each identity-versus-tail gap was at least g and remains strictly positive, so all identity pairs remain mutual top. The stable set, stable-pair graph, and A-optimal matching remain unchanged throughout that open ℓ_∞ ball. The three target margins are consequently at least $g/2$.

□

Property H.2 (Margins, information, and refinement)**.** *For every unique stable-market target,* $\kappa_\Psi(\theta)^2/(2d) \leq \Gamma^*_\Psi(\theta) \leq \kappa_\Psi(\theta)^2/2$*. If* $\Psi_1 \succeq \Psi_2$*, then* $\kappa_{\Psi_1} \leq \kappa_{\Psi_2}$*,* $\Gamma^*_{\Psi_1} \leq \Gamma^*_{\Psi_2}$*, and* $T^*_{\Psi_1} \geq T^*_{\Psi_2}$*.*

Proof of Property H.2. Let $a = \Psi(\theta)$ be the unique correct answer. We first justify $\kappa_\Psi(\theta) = \inf_{\lambda \in \mathrm{cl}(\mathrm{Alt}_a)} \|\theta - \lambda\|_\infty$. A segment from θ to any strict $\lambda \in \mathrm{Alt}_a$ has a first point at which the chamber label ceases to be a; that point lies in $\mathfrak{D}_\Psi$, so the left side is no larger than the right. Conversely, if $x \in \mathfrak{D}_\Psi$, arbitrarily small neighbourhoods of x contain strict chambers with different labels. If neither label is a, the segment from θ to either chamber first crosses an a-changing face no farther from θ than x; moving an arbitrarily small distance beyond that face gives a point of Alt_a. Taking infima proves the equality. For the uniform allocation $w_e = 1/d$, every $\lambda \in \mathrm{cl}(\mathrm{Alt}_a)$ satisfies $I_w(\theta, \lambda) = \|\theta - \lambda\|_2^2/(2d) \geq \kappa_\Psi(\theta)^2/(2d)$, which proves the lower bound on Γ^*_Ψ.

For the reverse inequality, fix $\zeta > 0$ and choose $x \in \mathfrak{D}_\Psi$ with $\|x - \theta\|_\infty \leq \kappa_\Psi(\theta) + \zeta$. By the definition of the discriminant, there is a strict $\lambda \in \mathrm{Alt}_a$ arbitrarily close to x. Hence, for every allocation w, $I_w(\theta, \lambda) \leq \|\theta - \lambda\|_\infty^2/2$. Sending first $\lambda \to x$ and then $\zeta \downarrow 0$ gives $\inf_{\lambda \in \mathrm{cl}(\mathrm{Alt}_a)} I_w(\theta, \lambda) \leq \kappa_\Psi(\theta)^2/2$. Maximizing over w proves the upper bound.

Suppose now that $\Psi_1 \succeq \Psi_2$. If two chambers have different Ψ_2-values, they necessarily have different Ψ_1-values. Therefore $\mathfrak{D}_{\Psi_2} \subseteq \mathfrak{D}_{\Psi_1}$ and $\kappa_{\Psi_1} \leq \kappa_{\Psi_2}$. For their correct answers at θ, the target-changing alternatives satisfy $\mathrm{Alt}_{\Psi_2} \subseteq \mathrm{Alt}_{\Psi_1}$. Taking an infimum over the larger alternative set can only reduce the information rate, so $\Gamma^*_{\Psi_1} \leq \Gamma^*_{\Psi_2}$ and inversion yields $T^*_{\Psi_1} \geq T^*_{\Psi_2}$. □

Proof of Theorem 3.2. For one candidate pair p, an ℓ_∞ perturbation of radius r can increase either comparison gain by at most $2r$, because its candidate coordinate may move up by r and its current-partner coordinate down by r. Hence making both gains nonnegative requires $r \geq \max\{\Delta_p^A, \Delta_p^B\}/2$. This radius is sufficient: move the two coordinates of each negative endpoint comparison toward one another by half its deficit, leaving an already nonnegative side unchanged. Since instability is the union over candidate pairs, minimizing over p proves the margin formula.

Let κ denote this margin. Under the uniform allocation $w_e = 1/d$, every comparison harmonic weight equals $h(1/d, 1/d) = 1/(2d)$. For every p, at least one of its deficits equals or exceeds 2κ, so $I_p(w) \geq (2\kappa)^2/(4d) = \kappa^2/d$. Maximizing over allocations gives the lower rate bound. The upper rate bound is the unique-target inequality in Property H.2. □

H.1 A SMALL GAP CAN AFFECT ONLY ONE TARGET

In Table 4, only the displayed comparison of B-agent 1 changes. Choose its cardinal gap ε and every other adjacent gap at least one.

Proposition H.3 (Graph sensitivity without optimal-matching sensitivity)**.** *Profile* $+$ *has stable matchings* $(1, 2, 3)$ *and* $(3, 2, 1)$*; profile* $-$ *has only* $(1, 2, 3)$*. The* A*-optimal matching is* $(1, 2, 3)$ *in both. Consequently, at the* $+$ *profile* $\kappa_G \leq \varepsilon/2$ *while* $\kappa_{M^A} \geq 1/2$*. In this family,* $T^*_G/T^*_{M^A} = \Omega(\varepsilon^{-2})$*.*

Proof. There are six perfect matchings. The following table gives, for every nonstable matching, one explicit blocking pair; "stable" means that direct inspection of all six unmatched pairs finds none.

Table 4: The stable-pair graph changes while the A-optimal matching does not.

Side	agent	profile $+$	profile $-$
A	1	$1 \succ 2 \succ 3$	$1 \succ 2 \succ 3$
A	2	$1 \succ 2 \succ 3$	$1 \succ 2 \succ 3$
A	3	$3 \succ 1 \succ 2$	$3 \succ 1 \succ 2$
B	1	$3 \succ 1 \succ 2$	$1 \succ 3 \succ 2$
B	2	$2 \succ 1 \succ 3$	$2 \succ 1 \succ 3$
B	3	$2 \succ 1 \succ 3$	$2 \succ 1 \succ 3$

Table 5: The stable set is identical but its A-optimal element differs.

Side	agent	profile I	profile II
A	1	$2 \succ 1 \succ 3$	$3 \succ 1 \succ 2$
A	2	$2 \succ 3 \succ 1$	$2 \succ 1 \succ 3$
A	3	$2 \succ 1 \succ 3$	$2 \succ 1 \succ 3$
B	1	$3 \succ 2 \succ 1$	$1 \succ 2 \succ 3$
B	2	$3 \succ 2 \succ 1$	$1 \succ 3 \succ 2$
B	3	$3 \succ 1 \succ 2$	$3 \succ 2 \succ 1$

matching	profile $+$	profile $-$
$(1,2,3)$	stable	stable
$(1,3,2)$	$(2,2)$	$(2,2)$
$(2,1,3)$	$(1,1)$	$(1,1)$
$(2,3,1)$	$(2,2)$	$(2,2)$
$(3,1,2)$	$(2,2)$	$(2,2)$
$(3,2,1)$	stable	$(1,1)$

For example, in matching $(1,3,2)$, agent 2 is assigned to 3 but ranks 2 above 3, while partner 2 is assigned to 3 and ranks agent 2 above agent 3; hence $(2,2)$ blocks in both profiles. The other listed certificates are checked identically. Thus the stable sets are exactly those in the proposition.

Under A-proposing deferred acceptance, agents 1 and 2 first propose to partner 1, partner 1 retains agent 1 in both profiles, agent 2 then proposes to partner 2, and agent 3 proposes first to partner 3. No later rejection changes these tentative assignments, so the resulting matching is the identity in both profiles.

Moving the two utilities in the exchanged comparison of B-agent 1 by $\varepsilon/2$ toward one another reaches the tie face between the two displayed chambers. Since the stable-pair graph differs across that face, $\kappa_G \le \varepsilon/2$.

The A-optimal matching is the identity on both chambers incident to that tie face, so the relative interior of the face is not in the M^A discriminant. Every other adjacent comparison has cardinal gap at least one. Reaching any other face of the order arrangement in ℓ_∞ norm requires perturbation at least $1/2$. Any lower-dimensional face that can change M^A also lies in one of those other tie hyperplanes. Hence $\kappa_{M^A} \ge 1/2$.

For the stable-pair graph in Proposition H.3, the target margin is at most $\varepsilon/2$. The upper half of Property H.2 gives $\Gamma^*_G \le \varepsilon^2/8$ and hence $T^*_G \ge 8/\varepsilon^2$. The A-optimal target has margin at least $1/2$. With $d = 18$, the lower half of Property H.2 gives $\Gamma^*_{M^A} \ge 1/(8d)$ and hence $T^*_{M^A} \le 8d$. Therefore $T^*_G/T^*_{M^A} \ge 1/(d\varepsilon^2)$.

□

H.2 The stable set need not identify its proposer-optimal element

Proposition H.4 (Non-refinement of stable-market targets)**.** *Both profiles in Table 5 have stable set* $\{(1,3,2),(3,1,2)\}$ *and hence the same stable-pair graph. Their A-optimal matchings are* $(1,3,2)$

and $(3,1,2)$*, respectively. Therefore neither* $G \succeq M^A$ *nor* $M^A \succeq G$*, and the unordered stable set does not refine* M^A*.*

Proof. Exhausting the six perfect matchings gives the following blocker certificates.

matching	profile I	profile II
$(1,2,3)$	$(3,1)$	$(3,2)$
$(1,3,2)$	stable	stable
$(2,1,3)$	$(2,2)$	$(1,1)$
$(2,3,1)$	$(2,2)$	$(1,1)$
$(3,1,2)$	stable	stable
$(3,2,1)$	$(3,2)$	$(3,2)$

Every listed pair strictly prefers each other to its assigned partners in the corresponding profile. Direct inspection of the unmatched pairs of $(1,3,2)$ and $(3,1,2)$ finds no blocker in either profile. Hence these and only these two permutations are stable.

In profile I the first proposal round is $1 \to 2$, $2 \to 2$, and $3 \to 2$; partner 2 retains agent 3. Continuing the rejection chain produces $(1,3,2)$. In profile II, the analogous proposal and rejection sequence produces $(3,1,2)$. Each returned matching is one of the two verified stable outcomes and is therefore the A-optimal element of its profile.

Since the stable sets are identical, their unions of edges are identical, but their A-optimal elements differ. Therefore neither the stable-pair graph nor the unordered stable set determines the A-optimal element. Conversely, Proposition H.3 gives two profiles with the same A-optimal matching and different stable-pair graphs, so M^A does not determine G.

□

I Pilot robustness and finite confirmation

This comparison pays to estimate all endpoint signs and therefore has a stronger gap requirement than guard racing.

Lemma I.1 (Uniform robustness of the scalar audit game)**.** *Suppose an estimate preserves every gain sign and estimates each positive protection deficit, and each positive gain of a true blocker, within a factor* $1 \pm \epsilon$*. Choose its optimal stable design or its highest-rate blocker and that blocker's optimal design. The resulting* w^γ *has true answer-specific rate at least* $c_{\epsilon,\gamma}\Gamma_\theta$*. The fitted optimum lies between* $(1-\epsilon)^2\Gamma_\theta$ *and* $(1+\epsilon)^2\Gamma_\theta$*. If every absolute gain is at least* $g_{\min}$*, a uniform pilot with*

$$m \geq \frac{4}{\epsilon^2 g_{\min}^2} \log \frac{4n(n-1)}{\alpha}$$

ensures these hypotheses with probability at least $1-\alpha$*.*

Proof. Every stable pair rate and every blocker's invalidation rate is a sum or minimum of nonnegative harmonic terms multiplied by squared gains. Relative gap accuracy sandwiches each such function between $(1-\epsilon)^2$ and $(1+\epsilon)^2$ times its true value, uniformly over allocations and answers. Taking minima and maxima preserves this sandwich. Evaluating the fitted optimizer in the true game loses at most $[(1-\epsilon)/(1+\epsilon)]^2$; sign preservation ensures its selected blocker is valid. The uniform mixture loses at most another factor $1-\gamma$ by monotonicity and homogeneity. Each pilot gain has Gaussian error variance $2/m$; union-bounding $2n(n-1)$ two-sided tails gives failure probability at most $4n(n-1)\exp(-m\epsilon^2 g_{\min}^2/4)$. This proves the claim for both stable and unstable answers with one argument. □

For a star, let $\widehat{H}$ be the cost computed from its estimated deficits. Inversion gives $H/(1+\epsilon)^2 \leq \widehat{H} \leq H/(1-\epsilon)^2$; the general robustness bound therefore also covers the star formula.

Theorem I.2 (Finite pilot and confirmation cost)**.** *Assume $|g_p^s| \geq g_{\min} > 0$ for all p and s, and take the m-sample-per-coordinate pilot of Lemma I.1 at failure level α. From the pilot compute the stable plug-in design or the easiest empirical blocker design, and put $w^\gamma = (1-\gamma)\widehat{w} + \gamma \mathbf{1}/d$, where $0 < \gamma < 1$. In an independent confirmation phase, use nested checkpoints with $N_e(t) = \lceil t w_e^\gamma \rceil$ fresh samples from coordinate e, whose total is at most $t + d$. At a checkpoint, report* `stable` *if the empirical blocker set is empty and $Z_{\mathrm{st}} \geq \beta_{\mathrm{sc}}(t+d,\delta)$; if that set is nonempty, report the lexicographically first maximizer of Z_p over its members when $\max_{p \in \widehat{\texttt{block}}} Z_p \geq \beta_{\mathrm{sc}}(t+d,\delta)$, where $\widehat{\texttt{block}}$ is the empirical blocker set at that checkpoint. Define $c_{\epsilon,\gamma} = (1-\gamma)[(1-\epsilon)/(1+\epsilon)]^2$ and set $\Gamma_\theta = \Gamma_M^{\mathrm{st}}$ when M is stable and $\Gamma_\theta = (T_M^{\mathrm{un}})^{-1}$ otherwise. For any $0 < \eta, \zeta < 1$, if a checkpoint t satisfies*

$$t \geq \frac{8d}{\gamma \eta^2 g_{\min}^2} \log \frac{2d}{\zeta}, \qquad\qquad t(1-\eta)^2 c_{\epsilon,\gamma} \Gamma_\theta \geq \beta_{\mathrm{sc}}(t+d,\delta),$$

then the procedure has stopped by at most $dm + t + d$ observations with probability at least $1 - \alpha - \zeta$. For every pilot outcome it is δ-correct; hence it has stopped with a correct answer by that budget with probability at least $1 - \alpha - \zeta - \delta$. The leading confirmation cost is the oracle characteristic time inflated only by $[(1-\eta)^2 c_{\epsilon,\gamma}]^{-1}$, plus the displayed pilot and concentration transient.

Proof of Theorem I.2. Let $\mathcal{E}_{\mathrm{pil}}$ be the event that every pilot gain has the correct sign and relative error at most ϵ. The calculation in Lemma I.1 gives $\mathbb{P}(\mathcal{E}_{\mathrm{pil}}) \geq 1 - \alpha$. The same lemma gives true answer-specific rate at least $c_{\epsilon,\gamma}\Gamma_\theta$ on this event.

Condition on the complete pilot. The confirmation allocation is then fixed, the nested checkpoint schedule is deterministic, and its observations are independent Gaussian samples. At virtual checkpoint t, write $N_e = \lceil t w_e^\gamma \rceil$; these counts require at most $t + d$ fresh observations. If their actual total is S_t, then $S_t \leq t + d$, and the monotonicity of $\beta_{\mathrm{sc}}(\cdot, \delta)$ makes the threshold $\beta_{\mathrm{sc}}(t+d,\delta)$ at least as conservative as the anytime threshold $\beta_{\mathrm{sc}}(S_t, \delta)$. A union bound over coordinates gives, with conditional probability at least $1 - \zeta$,

$$|\widehat{\theta}_e - \theta_e| \leq \sqrt{\frac{2\log(2d/\zeta)}{N_e}} \leq \sqrt{\frac{2d\log(2d/\zeta)}{t\gamma}} \qquad (e = 1, \ldots, d).$$

Under the first displayed condition in the theorem, every empirical current-versus-candidate gain therefore differs from its true value by at most $\eta g_{\min}$. All signs are preserved and every nonzero gain or deficit is estimated within a factor $1 \pm \eta$.

The integer allocation dominates the fractional one coordinatewise: $N_e \geq t w_e^\gamma$. Since the harmonic mean is coordinatewise nondecreasing and positively homogeneous, every empirical projection cost at the confirmation checkpoint is at least $t(1-\eta)^2$ times the corresponding true per-sample cost under w^γ. In the stable case this holds simultaneously for all candidate blockers, so $Z_{\mathrm{st}} \geq t(1-\eta)^2 c_{\epsilon,\gamma}\Gamma_\theta$ and the empirical blocker set is empty. In the unstable case it holds for the pilot-selected true blocker on $\mathcal{E}_{\mathrm{pil}}$; that pair belongs to the empirical blocker set, so the maximum blocker statistic is at least the same quantity. The second displayed condition makes the relevant statistic at least $\beta_{\mathrm{sc}}(t+d,\delta)$, and the all-answer checkpoint rule has therefore stopped. Combining the pilot and confirmation events proves the $1 - \alpha - \zeta$ stopping claim.

For correctness, condition again on an arbitrary pilot realization. The frozen allocation and its nested integer schedule are predictable for the fresh confirmation sequence. If the rule reports an incorrect answer a at a checkpoint, the true parameter lies in the alternative set invalidating a, so the stopped statistic satisfies $Z_a \leq L_{S_t}(\theta)$. The mixture martingale argument of Lemma D.2 and the inequality $Z_a \leq L_{S_t}(\theta)$ for an incorrect answer therefore bounds the probability of any wrong report by δ, regardless of the pilot signs and regardless of the maximization over empirical blocker answers. Because $\gamma > 0$, every confirmation count diverges. The strong law makes the true answer's statistic grow linearly, while β_{sc} grows logarithmically, proving almost-sure stopping for every pilot realization. Removing the conditioning proves unconditional δ-correctness. A final union bound gives the correct-by-budget statement. The interpretation of the leading cost follows by solving the second displayed inequality for t and observing that $\beta_{\mathrm{sc}}(t+d,\delta) = \log(1/\delta) + O(d \log\log(1/\delta))$ for fixed instance and $\delta \downarrow 0$. □

Table 6: Median total observations on doubly protected stable instances. Uniform uses the same audit target. "All signs" is a deliberately harder diagnostic that learns both endpoint signs; it is not a competing auditor.

n	Redundant gap scale	Guard audit	Uniform audit	All signs
8	0.1	29,888	65,536	1,851,392
8	0.01	29,888	65,536	202,899,456
8	0.001	29,888	65,536	20,333,985,792
32	0.1	549,168	2,097,152	34,430,976
32	0.01	549,168	2,097,152	3,652,714,496
32	0.001	549,168	2,097,152	376,279,400,448

J Experimental protocols and complete comparisons

J.1 Local tracking implementation

The numerical implementation initializes with n cyclic matchings, exposing every edge once. It then deploys batches of 128 independent matchings from a Birkhoff mixture, retaining all returned bilateral observations. On empirically stable instances, a numerical SLSQP solve maximizes the minimum of the concave pair information functions subject to row and column exposure sums one. The code checks the row/column and epigraph residuals and falls back to uniform exposure after a failed solve. A small balancing correction used before decomposition is explicitly recorded. On unstable instances it selects a highest-information empirical blocker and uses the half–half swap mixture. The two tested unstable families have a unique true blocker; the general asymptotic theorem instead uses the stated sticky tie rule when several optimal witnesses coexist. Every design is mixed with uniform exposure of weight $t^{-1/4}$. Stable designs are recomputed at geometrically increasing times. Finite solver tolerance and batched updates are disclosed numerical choices, not replacements for the vanishing-tolerance oracle required by the asymptotic theorem.

The two boundaries are evaluated on the same continuing adaptive path, including observations after the first boundary would have stopped. Thus their difference isolates the stopping-boundary overhead for that path. Reported elapsed times include this continuation and numerical optimization; they are not the latency of the earlier local stop. Gaussian sufficient statistics for an edge observed k times are simulated as $k\theta + \sqrt{k}Z$ with $Z \sim \mathcal{N}(0,1)$, independently on the two sides. This is exactly the distribution of the observed sums, not a reduction in the counted number of deployments. All seeds, counts, solver diagnostics, and censoring flags are provided.

J.2 Scalar guards and matching guards

The 1800 epoch experiments use 30 seeds for each size, margin setting, truth status, and applicable rule. Stable instances have heterogeneous A-side rejections in $[0.5, 1.5]$, while all B-side rejections lie in $[\varepsilon, 1.2\varepsilon]$. The same A-side instances and random streams are used when ε changes. A second family flips one pair to create a blocker. All runs return the correct audit answer. Batch Gaussian sufficient statistics simulate the exact observation-count law; a reported count is a query count, not the number of scalar random numbers generated by the implementation.

At $n = 32$, scalar guards use $549{,}168$ median observations for every redundant-margin setting, versus $2{,}097{,}152$ for same-target uniform auditing. Resolving all signs is a harder diagnostic, not evidence for an optimal information constant.

Initialization costs are studied separately in Appendix J.3. Table 9 measures information rates rather than stopping times.

The deployment-round study uses the identity matching and four families. Redundant instances have A-deficits uniform on $[0.3, 2]$ and B-deficits 0.001. Mixed instances independently choose one rejecting side per pair, with deficit uniform on $[0.3, 2]$ and the other gain uniform on $[0.3, 1.5]$. Double instances use deficits from those two intervals on the two sides. Homogeneous instances have $g_p^A = -1$ and $g_p^B = 0.5$. It runs 320 stable-audit instances: four sizes, four families, two methods, and ten seeds. Both methods observe feasible matchings, reuse completion-edge observations, and

Table 7: Median total observations, including initialization. The same plug-in rule is compared with two explicitly different oracle initialization costs.

Rule	Stable	Unstable
Oracle 64	49,800	20,800
Oracle 1	38,700	7,200
Plug-in 64	60,560.5	21,550

apply the same stopping rule. Table 2 includes wall-clock time as well as deployments. Homogeneous instances have a known characteristic time and expose the finite-confidence overhead rather than suggesting that a logarithmic-factor theorem implies a ratio near one. All individual trials, including nontermination flags, are saved.

J.3 Oracle initialization

An initialization ablation keeps the likelihood stopping rule, market, and plug-in rule fixed. Each cell below uses 50 seeds at $n = 10$ and confidence 0.95. All 300 answers are correct. "Oracle 1" knows the true optimal allocation and takes only one initial sample per coordinate; "Oracle 64" shares the plug-in's full initialization.

The unstable plug-in/oracle ratio is 1.0361 with common initialization, but 2.9931 against Oracle 1. Thus common-pilot comparisons do not establish end-to-end near-oracle cost. The separate full two-sided design experiment in Table 9 measures information-rate retention, not stopping time. We retain its raw data and distinguish that diagnostic from both total-cost experiments.

J.4 Stable and unstable matching-feedback comparisons

A separate matching-feedback study runs 320 algorithm instances, with 20 independent seeds per setting and paired seeds across comparators, for $n \in \{4, 8, 16, 32\}$. Every deployment is feasible and both utilities are observed. At $n = 32$, median stable round counts are 23,488 for guards and 131,072 for uniform edge coverage, a ratio of 5.5804; unstable medians are 6,800 and 16,384. All 320 outputs are correct. These are finite synthetic checks with nominal error 0.05, not an empirical assertion of zero risk or of asymptotic optimality. The code draws exact Gaussian batch sums, so large sample counts do not require materializing every individual observation.

J.5 Pilot designs and two-sided stress tests

We evaluate the same anytime GLR certificate at confidence 0.95 under three sampling rules. The *oracle* uses the true information-optimal design. The *plug-in* rule first takes 64 samples from every utility coordinate, estimates all current-versus-candidate gains, solves the estimated star or blocker design, and mixes it with 1% uniform exploration; it never reads the true utilities. The *uniform* rule samples every coordinate equally. Stable instances use an identity matching singly protected by the A side; unstable instances reverse one pair into a blocker. We use 50 independent trials for each method, truth, and market size.

At $n = 10$, the stable plug-in median is 64,064, compared with 49,300 for the oracle and 385,300 for uniform auditing. On the unstable instance, the corresponding medians are 21,800, 20,800, and 979,300. Across these common-initialization settings the plug-in/oracle ratio is between 1.04 and 1.30. The separate initialization ablation in Appendix J.3 shows why this is not an end-to-end near-oracle guarantee. The rate behavior is consistent with Lemma I.1: a short pilot need only estimate the protection geometry well enough to preserve the max–min rate; the safety mixture protects against a bad pilot without supplying the main evidence.

The right panel is a direct falsification test for global minimum-gap accounting. The changing comparison is between two alternatives that are both below the current partner, so it cannot create a blocker and is absent from the audit discriminant. Its shrinking gap has no visible effect on either audit rule, yet complete ranking pays an increasing inverse-square cost. The scripts regenerate the trial-level CSV files, table, and vector figure.

Table 8: Median total samples at confidence 0.95. All 900 audit runs return the correct answer. All three rules share a 64-sample-per-coordinate initialization. The last column is a controlled, common-initialization comparison, not the total cost relative to a minimally initialized oracle.

n	truth	oracle	plug-in	uniform	plug-in/oracle
5	stable	7200	7700	21200	1.07
5	unstable	5200	5700	54200	1.10
8	stable	26192	31694	150692	1.21
8	unstable	13692	14192	388942	1.04
10	stable	49300	64064	385300	1.30
10	unstable	20800	21800	979300	1.05

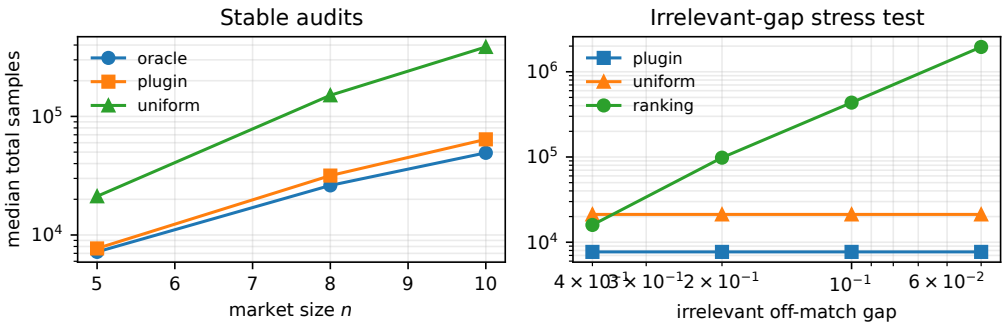


Figure 1: Left: stable-audit scaling. Right: an off-match comparison is moved toward a tie while all audit-relevant gaps stay fixed. The plug-in and uniform audit medians remain flat, whereas uniform complete ranking grows from 15,950 samples at gap 0.40 to 1,954,450 at gap 0.05. Curves report medians over 50 trials.

General two-sided stress test. The closed-form scaling family is singly protected, so we additionally generate ten heterogeneous stable markets at each size in which every off-matching pair is protected by both endpoints. For every market we solve the full concave max–min design, estimate all utilities from an independent pilot of 1024 samples per coordinate, resolve the estimated max–min program, and mix its optimizer with 1% uniform exploration. Table 9 reports the true worst-blocker information rate of each allocation divided by the true oracle rate; it evaluates design quality directly, before any stopping-bound overhead.

All pilot signs are correct in the thirty markets. The plug-in design retains median oracle rate between 0.871 and 0.900, and its worst observed ratio is 0.839; uniform sampling falls from 0.592 at $n = 4$ to 0.399 at $n = 8$. This test is outside the star formula and uses both endpoint certificates for every candidate pair. It therefore checks the scope of Theorem I.2: learning the full two-sided geometry preserves most of the max–min information advantage, rather than merely recovering a closed-form special case.

K Additional related work

Classical theory establishes deferred acceptance, lattice structure, rotations, enumeration, and polyhedral descriptions (Gale & Shapley, 1962; Knuth, 1976; Irving & Leather, 1986; Gusfield, 1987; Gusfield & Irving, 1989; Feder, 1992; Vande Vate, 1989; Rothblum, 1992; Teo & Sethuraman, 1998). Stable roommates, incomplete lists, ties, weak stability, and super stability lead to different existence and complexity phenomena (Irving, 1985; 1994; Manlove et al., 2002; Basu, 2025). Preference elicitation and uncertain-preference work asks which deterministic comparisons, interviews, or partial

Table 9: Rate retention on heterogeneous doubly protected audits. Each row contains ten independently generated markets; p10 is the empirical 10th percentile, and higher is better.

n	plug-in p10	plug-in median	plug-in min	uniform median	signs correct
4	0.851	0.900	0.839	0.592	100.0%
6	0.846	0.878	0.840	0.485	100.0%
8	0.850	0.871	0.839	0.399	100.0%

Table 10: The closest learning formulations differ in the statistical object that must be certified.

Line of work	Required output	Distinction from this paper
Stable regret	one repeatedly played stable outcome	cumulative loss; no fixed-confidence equilibrium certificate
Stable identification	one stable or optimal stable matching	does not certify a prescribed deployment or the full stable structure
Fair stable optimization	one optimizer over a known stable set	preferences are observed; statistical acquisition is absent
Noisy ranking	complete preference orders	learns comparisons that may lie outside the target discriminant
This paper	audit answer or specified stable-market target	target-specific alternatives, exact audit design, and protection sharing

orders suffice for a stable output (Immorlica & Mahdian, 2005; Drummond & Boutilier, 2013; Rastegari et al., 2013; Genc et al., 2017; Aziz et al., 2020). We study adaptive noisy observations and equilibrium-structure targets rather than noiseless elicitation of one outcome. Matching bandits study cumulative regret and coordination (Liu et al., 2020; 2021; Cen & Shah, 2021; Sankararaman et al., 2021; Basu et al., 2021; Jagadeesan et al., 2023; Kong & Li, 2023; Kong et al., 2024; 2025; Wu et al., 2026). The closest fixed-confidence work identifies one stable or optimal stable matching (Hosseini et al., 2024; Pagare et al., 2025; Athanasopoulos et al., 2026). In particular, Athanasopoulos et al. (2026) learn an optimal stable matching when both sides are unknown under semi-bandit feedback. Their output is a market-selected optimum; our primary output certifies a prescribed deployment, has a universal stable answer and multiple unstable witnesses, and admits an exact asymptotic audit design. Methodologically, the information lower bound and sticky answer selection build on fixed-confidence pure exploration (Even-Dar et al., 2006; Audibert & Bubeck, 2010; Kalyanakrishnan et al., 2012; Karnin et al., 2013; Bubeck et al., 2013; Chen et al., 2014; Jamieson et al., 2014; Garivier & Kaufmann, 2016; Kaufmann et al., 2016; Chen et al., 2017; Degenne & Koolen, 2019; Fiez et al., 2019; Kaufmann & Koolen, 2021). Poiani et al. (2026) give finite-confidence guarantees for generic Track-and-Stop and Sticky Track-and-Stop. Our contribution is matching specific: the bilateral audit information game, a polynomial blocker oracle, and a pilot-free guard bound that avoids learning unnecessary endpoint signs. The Gaussian star optimizer is an inherited best-arm subproblem, and generic tracking is not claimed as a new algorithmic principle. Noisy sorting and active ranking recover complete orders (Feige et al., 1994; Braverman & Mossel, 2008; Jamieson & Nowak, 2011; Heckel et al., 2019); our discriminant shows why complete ranking can be statistically unnecessary.

L Reproducibility details

The statistical-certificate calculations are generated by `run_statistical_design_certificates.py`; their complete records are in `data/statistical_design`. Repeated looks within a path are not counted as independent trials.

The numerical design checks are generated by `run_design_certificates.py`; raw summaries are in `data/design_certificates`. These checks supplement, rather than replace, the analytic proofs.

The local-boundary experiment is reproduced by `run_local_tracking.py --trials 30`; its 240 paths retain both stops, solver diagnostics, and full-path timings.

The matching-round extension is reproduced by `run_guard_rounds.py`; raw trials include deployment count, scalar observation count, scheduling batches, time, and correctness. Theorem 4.1 applies to this implemented guard; the separate Birkhoff tracker retains only its stated asymptotic guarantee.

Complete arguments are provided in the main text and appendices. The script `run_guard_racing.py` regenerates all 1800 adaptive-guard, uniform-audit, and all-sign diagnostic runs, including the exact Gaussian batch counts. The script `check_oracle_initialization.py` reproduces the 300 total-cost oracle and plug-in comparisons. The script `run_matching_feedback.py` generates 320 matching-feasible bilateral-feedback runs and exact-design checks. Information-design experiments and `verify_theory.py` are also retained. Seeds, individual outcomes, settings, and numerical tolerances are supplied; finite checks do not replace proofs.