WHEN DO LLM AGENTS SETTLE? MEMORY-MEDIATED POTENTIAL GAMES

Anonymous authors
Paper under double-blind review

ABSTRACT

Multi-agent systems increasingly revise persistent shared state such as memories, workboards, and partial solutions. Each worker owns the part it may rewrite and writes only what it judges better than the incumbent. Those judgments are imperfect, and the record used to evaluate a candidate may be stale. When must such revision settle? We study this in finite exact-potential games, where a worker's report of a candidate's improvement may differ from the truth by at most β and a revision is accepted only when the reported gain exceeds a margin δ. We characterize the settling boundary: universal termination holds if and only if $\delta \geq \beta$. Above the boundary, every accepted revision makes true progress of more than $\delta - \beta$, so there is a finite move bound, and an exhaustive quiet scan gives a $(\delta + \beta)$-approximate Nash guarantee. At equality, termination survives without a spread-based rate; below the boundary, even stationary reports can cycle with one player. Representation matters too, since reports generated by one fixed scalar score terminate for every positive margin in that single-player case. Within the scalar class, let $n \geq 2$ workers update one at a time with pointwise scalar error at most η. Each reads its own coordinate fresh and others coordinatewise, possibly from different rounds at most $q \in \{1, \ldots, n-1\}$ rounds old. If changing one opponent's action changes a worker's true cost by at most $L > 0$ at every fixed own action, universal termination holds exactly when $\delta \geq 2\eta + L \min\{2q, n-1\}$, even with partial scans or unfair schedules. Matching finite games cycle below it. Potential descent handles short delays; long delays need an argument that no worker returns to an action it left. Further results account for simultaneous writes, certify recorded paths and fresh endpoints, and separate a quiet scan from absorption under continued stochastic evaluation, where ties can move infinitely often almost surely. An appendix applies record-level checks to one language model's comparisons at fixed profiles.

1 INTRODUCTION

Language-model agent systems increasingly coordinate through persistent shared state: task lists, tagged notes, partial tables, memories, and files that one worker writes and the next reads (Huang et al., 2026; Chen et al., 2026). Workers read what others wrote, propose revisions, and commit changes that affect later decisions. Does the rewriting stop, and if it stops, what does the resting state guarantee?

If every worker knew the true value of each unilateral revision, finite exact-potential games would answer the first classically: every strict improvement moves a common potential in one direction, so improvement paths terminate (Monderer & Shapley, 1996). Agent systems instead act on estimated comparisons. A worker reports how much better a candidate is than the incumbent; that report can be wrong by at most β, and the candidate is written only when the reported gain exceeds an acceptance margin δ. These reports need not come from one latent scalar score, and the record a worker reads may be stale.

The assumptions are ours rather than observations. We require a finite exact potential game, with true costs and legal deviations fixed throughout each episode, and revisions confined to owned coordinates. Exact potentiality is a structural hypothesis about the class we study, not a property established for arbitrary agent teams. Shared-workboard architectures motivate the study of evaluation error and

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.

read freshness (Huang et al., 2026), as does the out-of-sync behavior measured by Guo et al. (2025); they do not establish the game or the guarantees analyzed here.

Our first result characterizes the settling boundary exactly: universal termination over bounded direct comparison fields holds if and only if $\delta \geq \beta$. Above the boundary every accepted revision reduces the mover's true cost by more than $\delta - \beta$, giving a finite move bound and, after an exhaustive terminal scan, a $(\delta + \beta)$-approximate Nash certificate. At equality accepted revisions still strictly decrease the potential, so termination survives without a spread-over-margin rate. Below it even a one-player game with stationary reports can cycle forever. Representation matters as well: arbitrary direct reports may contain cyclic inconsistencies, whereas comparisons from a single scalar evaluation obey global cycle constraints. With fresh reads and one player using stationary scalar-representable reports, the exact asymptotic ceiling on the ratio of moves to that move bound, as the potential range grows with δ, β fixed, falls from one to $(\delta - \beta)/\delta$, strictly below one when $\delta > \beta > 0$ (Propositions L.1 and L.2).

Within that scalar class, our main technical result gives an exact delay-dependent threshold. Theorem H.7 treats stale scalar evaluations with one worker updating per round: if pointwise scalar error is at most η, a worker's own coordinate is fresh and others are read coordinatewise at most q rounds old, and changing one opponent's action changes the evaluating worker's true cost by at most L, then universal termination holds exactly when $\delta \geq 2\eta + L \min\{2q, n-1\}$. The delay dependence grows and then saturates, and the regimes are proved differently: potential descent supplies the guarantee at short delays, while the saturated regime requires an argument preventing return to previously left actions. The threshold is therefore not the direct bound with a larger error budget substituted in. One family of finite games covers both regimes, using genuine past coordinate values and one stationary scalar score per player. Each delay guarantee covers every qualifying selection under its protocol, and each matching lower bound exhibits one fair cycling trajectory rather than asserting that every schedule cycles.

A second boundary appears when evaluation continues after a quiet round, since a revision rejected once may be accepted later. Under persistent independent scalar resampling, whenever the noise can overcome the margin, a tied two-action example moves infinitely often almost surely, and for uniform noise, two-player exact-potential games with $O(H^2)$ profiles, $H + 1$ of them reachable, and an H-independent floor on positive reachable transition probabilities can absorb almost surely yet take expected time exponential in H, which we determine up to a constant factor.

These results separate three notions that are easy to conflate: making genuine progress, certifying a stable endpoint, and reaching an absorbing state. That separation organizes our extensions to restricted proposals, heterogeneous workers, stale reads, and concurrent writes.

To connect the model to shared-agent architectures we give a workboard coverage game in which workers select tag bundles and receive marginal coverage credit. Its potential is Rosenthal's specialized to marginal rewards, and coincides with total coverage (Rosenthal, 1973); the connection is structural rather than a new construction. Bilò et al. (2025) study the complexity and approximability of minimizing Rosenthal's potential in monotone congestion games; our focus is how imperfect comparisons affect the dynamics and their terminal guarantees.

Our main contributions are:

- **An exact settling boundary.** Universal termination under bounded direct reports holds if and only if $\delta \geq \beta$, including equality and a stationary sub-threshold cycling witness.
- **An exact delay-dependent threshold.** For coordinatewise-stale scalar evaluations with fresh own coordinates and singleton updates, $\delta \geq 2\eta + L \min\{2q, n-1\}$ is exact, combining potential descent at short delays with a no-return argument in the saturated regime and a matching construction for every player count and delay.
- **Two further separations.** For one player with stationary scalar-representable reports, the exact asymptotic normalized ceiling is $(\delta - \beta)/\delta$, strictly below the general constant one when $\delta > \beta > 0$. Under persistent i.i.d. scalar resampling with $\delta < 2\eta$, ties can recur forever, and even $O(H^2)$-profile games with an H-independent floor on positive reachable transition probabilities can absorb almost surely in expected time exponential in H.
- **A path to shared workboards.** We instantiate the framework through a Rosenthal-potential workboard game and show how proposal coverage, heterogeneous errors, stale reads, and concurrent writes alter progress guarantees or endpoint certification.

Related work. Potential games, finite improvement paths, and range-over-margin counting are classical (Monderer & Shapley, 1996; Rosenthal, 1973). Multiplicative improvement margins buy convergence speed (Chien & Sinclair, 2011; Awerbuch et al., 2008). Adversarial time-varying multiplicative cost perturbations are studied by Balcan et al. (2013); their welfare analysis differs from the additive comparison thresholds here. Near-potential dynamics may cycle while approaching approximate equilibrium sets (Candogan et al., 2013). Other models concern probabilistic response mistakes (Ferraioli & Penna, 2015), logit response (Penna, 2018; Blume, 1993; Alós-Ferrer & Netzer, 2010), or pointwise utility noise in log-linear learning (Maddux et al., 2024). Lei & Shanbhag (2020) allow bounded coordinatewise delays in continuous potential games, with proximal responses and controlled optimization errors; our finite-action threshold instead permits persistent bounded report error. Deliberately stale decisions have a history in nonatomic routing: Fischer & Vöcking (2005) refresh a bulletin board of edge latencies every T units of time. In the fluid limit, for every $T > 0$ some two-link instance cycles under unmargined better response, and damping the migration rate restores convergence to a Wardrop equilibrium when the damping is tied to the refresh period, the path length and the latency slope, every path is sampled with positive probability, and latencies are strictly increasing. A deterministic acceptance margin acts as such a damping but falls outside their theorem. With unit total flow and unit revision rate per player, applying the calculation in their Lemma 2 and the proof of Proposition 1 to this cutoff gives a per-phase potential bound when the margin is at least four times the product of path length, latency slope and refresh period. Summed over phases, that bound gives finite total migration, so the flow converges, with an approximate Wardrop limit under uniform sampling, but it gives neither finite termination nor an exact equilibrium. We give exact thresholds and finite termination in atomic finite games, for every admissible schedule and report-error realization. A tolerance also appears in the payoff-based dynamics of Marden et al. (2009), screening out apparent improvements as ours does; there the estimation error is a design parameter, controlled by the exploration rate and phase length, and the guarantee is probabilistic, whereas ours is an adversarial report error under a fixed bound and the guarantee is pathwise. Nisan et al. (2008) study asynchronous best reply without a margin or report error: players reply exactly to what update messages tell them, and the difficulty comes from the schedule and from messages that may be delayed or lost.

Closest on the convergence calculation is Kearns et al. (2015), whose differentially private mediator simulates an α-noisy best response: an additive margin α applied to reported costs carrying pointwise error at most Δ. For congestion games with losses in $[0, 1]$, the proof of their Lemma 2 derives a move bound of their potential's maximum over $\alpha - 2\Delta$, relaxed to a stated bound of order (resources $\times$ players)$/\alpha$, and the proof of their Theorem 6, whose statement is a privacy utility bound, shows the improvement remaining at termination is at most $\alpha + 2\Delta$. We claim neither sufficient formula: both are prior art. Their analysis does not study arbitrary direct comparison fields, the exact boundary, or the stationary separation proved here. Our general reports need not be jointly representable by scalar costs; when they are, the model reduces after recentering to the pointwise-error interpretation with $\beta = 2\eta$, and the flow and cycle conditions behind that distinction are classical (Candogan et al., 2011; Lim, 2020). The scalar one-player ceiling likewise uses ordinary descent in the reported score. The exact delay threshold additionally needs a no-return argument and matching constructions. It concerns only the stated acceptance rule.

Language-model agents in games. Strategic behavior is measured in game benchmarks and repeated interactions (Duan et al., 2024; Akata et al., 2025; Tewolde et al., 2026; Mei et al., 2024); equilibrium also guides reasoning and debate (Zhu, 2025; Yi et al., 2025). See Sun et al. (2025) for a survey and Li et al. (2026) for equilibrium-algorithm discovery.

Closer to shared-record updates, Rodriguez (2026) analyzes artifact patches through actual local improvement and bounded coupling; EquiMem models strategic memory contribution and auditing (Meng et al., 2026). Chen et al. (2026) study approximate-potential worker coordination and grounded verifier-gated reflection. Here we isolate a different source of approximation: the underlying game is exact, while the error lies in the worker's reported improvement. Wu et al. (2026) analyze probabilistic stopping and candidate retention with noisy verifiers. These guarantees concern actual local improvement, memory calibration, or probabilistic verifier decisions. Our scalar result fixes the update rule and characterizes its exact delay-dependent margin.

Empirical nontransitivity of LLM judgments (Xu et al., 2025) motivates allowing direct comparisons; it does not establish a uniform cardinal error bound. Our guarantees are conditional on that bound and on the specified game.

2 MODEL

A finite cost game has players $N = \{1, \dots, n\}$, nonempty finite strategy sets S_i, profiles $S = \prod_i S_i$, and costs $c_i : S \to \mathbb{R}_{\geq 0}$; Remark B.2 gives a conditional expected-loss interpretation. A function $\Phi : S \to \mathbb{R}$ is an *exact potential* when every unilateral deviation $s_i \to a$ satisfies $\Phi(a, s_{-i}) - \Phi(s) = c_i(a, s_{-i}) - c_i(s)$. Its spread is $\Delta_\Phi = \max_s \Phi(s) - \min_s \Phi(s)$. A profile s is an ε-Nash equilibrium (NE) if $c_i(s) \leq c_i(a, s_{-i}) + \varepsilon$ for every i and $a \in S_i$.

Definition 2.1 (Direct noisy better response)**.** For a profile s and candidate $a \neq s_i$, write the true cost improvement as $\Delta_i(s, a) = c_i(s) - c_i(a, s_{-i})$. Fix $\beta \geq 0$ and $\delta > 0$. At comparison occurrence u, player i directly reports $\widehat{\Delta}_i^u(s, a)$. The report field is β-accurate when $|\widehat{\Delta}_i^u(s, a) - \Delta_i(s, a)| \leq \beta$ for every i, u, s, and $a \neq s_i$. A deviation is accepted when $\widehat{\Delta}_i^u(s, a) > \delta$. An asynchronous trajectory records only accepted moves; its tth move is generated at some occurrence u_t. It is maximal if it is infinite or is accompanied at its endpoint by an exhaustive scan that evaluates and rejects every nonincumbent unilateral candidate, each at its own occurrence $\tau(i, a)$. The direct field is *stationary* when each ordered-pair report is independent of its occurrence index.

The exclusion of the incumbent is substantive: a self-loop is not a move. The occurrence index permits arbitrary time variation. Direct reports for opposite orientations are not assumed to be negatives of one another.

Definition 2.2 (Scalar-representable subclass)**.** For fixed i, u, s_{-i}, call $\{(a, s_{-i}) : a \in S_i\}$ the *fiber*; a direct field is *complete* on it if it reports every ordered pair of distinct actions. A complete field is scalar-representable if some $\hat{c}_i^u$ satisfies $\widehat{\Delta}_i^u((a, s_{-i}), b) = \hat{c}_i^u(a, s_{-i}) - \hat{c}_i^u(b, s_{-i})$ for every ordered pair of distinct actions $a, b \in S_i$. It is scalar-stationary when such a representative can be chosen independently of u.

Scalar representability is equivalent to zero report sum on every directed closed walk of the fiber. Under the full β bound, the residual $\hat{c}_i^u - c_i$ of a representative has diameter at most β on the fiber and can be recentered to pointwise error at most $\beta/2$. Thus pointwise $|\hat{c}_i^u - c_i| \leq \eta$ implies the subclass with $\beta = 2\eta$, and the converse holds after a per-fiber shift. Completeness is essential: residuals can accumulate along a partial path. Proposition B.3 gives the proof. The general direct model is strictly broader and retains neither this recentering nor every scalar-only stationary result.

The terminal guarantee of Theorem 4.2 quantifies over every $a \in S_i$. A worker that samples candidate revisions does not choose from all of S_i, so we state separately what a restricted process certifies.

Proposition 2.3 (Proposal-relative guarantee)**.** *Suppose occurrence u exposes only a proposal set $P_i^u(s) \subseteq S_i$ containing the incumbent. Fix declared scan sets $Q = (Q_i)_{i \in N}$, $Q_i \subseteq S_i$, and call a restricted trajectory Q-maximal if it is infinite or its endpoint is accompanied by a finite scan that evaluates and rejects every nonincumbent element of each Q_i. Under Definition 2.1 with $\delta > \beta$, every Q-maximal restricted trajectory is finite and satisfies the move bound of Theorem 4.2 unchanged, and its terminal profile s^T satisfies $c_i(s^T) \leq c_i(a, s_{-i}^T) + \delta + \beta$ for every i and $a \in Q_i$. Exhaustiveness, $Q_i = S_i$ for every i, upgrades this to a $(\delta + \beta)$-Nash equilibrium. Without it no such bound holds: the true improvement available at s^T can be as large as Δ_Φ. A given run may of course stop at an equilibrium anyway; what fails is the guarantee, not every instance.*

Termination is therefore robust to restricted proposals, and stability is not. The appendix proves this, gives the witness, and states the state-conditional coverage needed for persistent proposals to recover an exhaustive scan. At a fixed profile, a quiet scan rejects all nonincumbent candidates in Q_i; it certifies $(\delta + \beta)$-stability only within these sets.

3 A WORKBOARD GAME WITH AN EXACT POTENTIAL

Let $\mathcal{T}$ be a finite tag set, with tags $\tau \in \mathcal{T}$. Worker i chooses an admissible bundle $s_i \in S_i \subseteq 2^{\mathcal{T}}$. Write $\ell_\tau(s) = |\{i : \tau \in s_i\}|$. Each tag has a nondecreasing concave coverage function

$g_\tau : \{0, \dots, n\} \to \mathbb{R}_{\geq 0}$ with $g_\tau(0) = 0$. Its marginal coverage is $r_\tau(k) = g_\tau(k) - g_\tau(k-1)$, and worker i receives

$$u_i(s) = \sum_{\tau \in s_i} r_\tau(\ell_\tau(s)), \qquad W(s) = \sum_{\tau \in \mathcal{T}} g_\tau(\ell_\tau(s)).$$

This is Rosenthal's congestion-game construction in the payoff convention (Rosenthal, 1973): tags are the resources, a bundle is a strategy, and what a tag pays depends only on how many workers hold it. His potential adds up, for each resource, the rewards at loads $1, \dots, \ell_\tau(s)$; taking r_τ to be the increments of g_τ collapses that sum to W. It is also a marginal-contribution utility design of the kind studied in distributed welfare games (Marden & Wierman, 2013; Vetta, 2002; Gopalakrishnan et al., 2014). When every worker draws one tag from a common admissible set, set $c_\tau(0) = 0$ and $c_\tau(k) = C - r_\tau(k)$ for $1 \leq k \leq n$, with a common $C > \max_{\tau,\, 1 \leq k \leq n} r_\tau(k)$. This turns the game into a symmetric singleton congestion game whose costs are positive at positive loads and nondecreasing, matching the convention of the cited source, and whose Rosenthal potential is $nC - W$. All its pure Nash equilibria then share one potential value (Bilò et al., 2025, Proposition 1), and since a potential minimizer is itself an equilibrium, every pure equilibrium is coverage-optimal. Neither the general workboard nor the dynamics below assume that structure, and the statement covers exact equilibria only; Proposition 5.2 treats ε-equilibria.

Proposition 3.1 (Workboard potential)**.** *The workboard coverage game has exact utility potential* $\Psi(s) = \sum_{\tau \in \mathcal{T}} \sum_{k=1}^{\ell_\tau(s)} r_\tau(k) = W(s)$. *Consequently, every strict unilateral utility improvement strictly increases global coverage, and every maximizer of* W *is a pure Nash equilibrium.*

Proof. The inner sum telescopes to $g_\tau(\ell_\tau(s))$. For a deviation $s_i \to s_i'$, a newly added tag τ contributes $r_\tau(\ell_\tau + 1)$ to both the utility and potential differences; a dropped tag contributes $-r_\tau(\ell_\tau)$ to both; common tags cancel. Thus $u_i(s_i', s_{-i}) - u_i(s) = \Psi(s_i', s_{-i}) - \Psi(s)$. The two consequences follow immediately. □

This proposition formalizes only this workboard. Concavity supplies diminishing marginal credit and the congestion reading; it is not needed for the telescoping identity itself. In the cost convention, use $c_i' = K_i - u_i$ for any $K_i \geq \max_s u_i(s)$. This makes costs nonnegative, converts utility improvements into cost decreases, and gives the exact cost potential $\Phi_{\mathrm{wb}} = -\Psi = -W$; negation preserves the potential spread.

The displayed marginal-credit rule is one sufficient construction. A second question is when an already specified local utility is reliably aligned with a global objective. The quantifier in the following statement is essential.

Proposition 3.2 (Robust local-to-global alignment)**.** *Fix a player* i*, finite strategy sets, and the utilities* $(w_j)_{j \neq i}$ *of the other players. Write* $\Omega_{-i} = \sum_{j \neq i} w_j$ *for their aggregate, so that* $\Omega = w_i + \Omega_{-i}$ *is the total utility;* Ω *need not be the workboard coverage* W *of Proposition 3.1. The following are equivalent:*

1. *for every utility realization* w_i*, every strict unilateral* w_i*-improvement by player* i *strictly improves* Ω*;*
2. Ω_{-i} *is independent of player* i*'s action.*

The proof is in Proposition L.3.

For a single fixed w_i the reverse implication is false, since local improvements may be absent or may fortuitously dominate an externality. The proposition is a robust class statement, closely related to the classical four-cycle characterization of exact potentials (Monderer & Shapley, 1996, Corollary 2.9). If utilities can instead be designed, the marginal-contribution rule $w_i^{\mathrm{MC}}(s) = J(s) - J(s_i^0, s_{-i})$ for an objective J and baseline action s_i^0 makes every unilateral utility difference equal the corresponding difference in J; taking $J = W$ reproduces the unilateral differences of the workboard credits. The workboard credits, robust externality-freeness, and marginal-contribution redesign are three different ways to obtain alignment; none follows from the words "additive utility" alone.

What the abstraction covers. On a common board, each worker writes only to its own tagged slot (Huang et al., 2026); in other systems workers patch a shared artifact (Rodriguez, 2026) or update

a shared memory (Meng et al., 2026), and in collaborative coding others change the code base an agent is working on (Guo et al., 2025). Our formal model is narrower. Player i picks from a fixed finite bundle set S_i, and an accepted move changes only coordinate i, so the model covers bounded memory split into owned coordinates; free-form text or code edits and multi-coordinate overwrites fall outside it. File locking is compatible with sequential writes, but it does not make workers evaluate the displayed marginal credit or make one worker's utility difference equal the change in the task objective. Both are premises of our model, not consequences of these architectures. The model therefore keeps incumbent, candidate, local values, report and transition apart, so a mismatch can be traced to proposal, evaluation or update rather than blamed on the potential model.

4 THE THRESHOLD AND TERMINAL GUARANTEE

Lemma 4.1 (Soundness of an accepted move)**.** *Every accepted move from s^t to s^{t+1} satisfies*

$$c_i(s^t) - c_i(s^{t+1}) > \delta - \beta, \qquad \Phi(s^t) - \Phi(s^{t+1}) > \delta - \beta.$$

Proof. Let u_t be the occurrence that generates the tth accepted move, and let $e_t = \widehat{\Delta}_i^{u_t}(s^t, s_i^{t+1}) - \Delta_i(s^t, s_i^{t+1})$. Acceptance and $e_t \le \beta$ give $\Delta_i = \widehat{\Delta}_i^{u_t} - e_t > \delta - \beta$. Exact potentiality identifies this improvement with the potential drop. □

Theorem 4.2 (Convergence and terminal error)**.** *Let G be a finite exact potential game and let $\delta > \beta$. Every maximal asynchronous trajectory terminates. If it makes m moves, then*

$$m \le \max\left\{0, \left\lceil \frac{\Delta_\Phi}{\delta - \beta} \right\rceil - 1\right\} \le \frac{\Delta_\Phi}{\delta - \beta},$$

and $m < \Delta_\Phi/(\delta - \beta)$ holds strictly whenever $m \ge 1$. Every terminal profile is a $(\delta + \beta)$-Nash equilibrium.

Proof. Every finite prefix of $m \ge 1$ moves satisfies, by Lemma 4.1, $\Delta_\Phi \ge \Phi(s^0) - \Phi(s^m) > m(\delta - \beta)$. Since this bounds all finite-prefix lengths uniformly, no trajectory is infinite, and integrality yields the stated integer bound. At a terminal profile s^T, every scanned deviation has $\widehat{\Delta}_i^{\tau(i,a)}(s^T, a) \le \delta$ at its own scan time $\tau(i, a)$. Its error is at least $-\beta$, so the true improvement is at most $\delta + \beta$. □

The bound counts accepted moves, not computation: one explicit scan costs $\sum_i(|S_i|-1)$ comparisons, and Δ_Φ may be exponential in a succinct encoding, so this is not a polynomial running-time guarantee.

Table 1 (Appendix K) summarizes the three margin regimes.

At equality, Lemma 4.1 still gives strict potential decrease. No profile can repeat, so a maximal trajectory has at most $|S| - 1$ moves. The terminal calculation in Theorem 4.2 remains valid and gives radius 2β, but $\delta - \beta = 0$ supplies no positive denominator.

Below the threshold the obstruction already occurs with stationary reports. One player has actions A, B, both true costs zero, and $\Phi \equiv 0$. Set the two ordered reports $\widehat{\Delta}(A, B) = \widehat{\Delta}(B, A) = \beta$. Both are admissible and accepted when $0 < \delta < \beta$, so the path alternates forever. Their sum is 2β, so this stationary field is not scalar-representable. Thus, over the admitted range $\delta > 0$, universal termination across finite exact-potential games and admissible reports holds exactly when $\delta \ge \beta$. The corner $\delta = \beta = 0$ is outside Definition 2.1.

For $\delta > \beta$ and $\Delta_\Phi > 0$, define the normalized rate as $m(\delta - \beta)/\Delta_\Phi$. A stationary direct staircase approaches normalized rate one without attaining it (Proposition L.1); for one player with stationary scalar-representable reports the exact asymptotic ceiling, as $\Delta_\Phi \to \infty$ with δ, β fixed, is instead $(\delta - \beta)/\delta$, strictly below one when $\delta > \beta > 0$ (Proposition L.2). So the two classes in Table 2 (Appendix K) have the same sufficient convergence bounds but not the same stationary behavior.

5 HETEROGENEOUS WORKERS

Mixed-model systems need not share one error budget or one acceptance margin. Give player i parameters $\beta_i \ge 0$ and $\delta_i > 0$, require its direct comparisons to be β_i-accurate in the sense of

Definition 2.1, and accept its move only when its reported improvement exceeds δ_i. A profile is a vector-$\boldsymbol{\varepsilon}$ equilibrium when player i can improve its true cost by at most ε_i.

Theorem 5.1 (Player-specific parameters)**.** *Suppose $\delta_i > \beta_i$ for every player and set $\alpha = \min_i(\delta_i - \beta_i) > 0$. Every maximal heterogeneous trajectory terminates. If at least one move occurs and m_i is the number of moves by player i, then*

$$\sum_i m_i(\delta_i - \beta_i) < \Delta_\Phi, \qquad \sum_i m_i \le \left\lceil \frac{\Delta_\Phi}{\alpha} \right\rceil - 1.$$

In all cases, $\sum_i m_i \le \max\{0, \lceil \Delta_\Phi/\alpha \rceil - 1\}$. Every terminal profile is a vector-$\boldsymbol{\varepsilon}$ equilibrium with $\varepsilon_i = \delta_i + \beta_i$.

This is the per-player constant special case of Proposition D.1, whose appendix proof also records the zero-move and sequential-scan cases.

The threshold must hold for each player, rather than only on average. If every $\delta_i \ge \beta_i$, each accepted move still strictly decreases the potential, so finiteness gives universal termination, though equality for any player removes the uniform spread rate. Conversely, if $0 < \delta_j < \beta_j$ for one player j, the stationary direct two-cycle of Proposition C.4 embeds while all other players remain at strict independent equilibria. Scalar stationarity is different: a lone mover always terminates by Proposition L.2. Both are stationary, so the difference is integrability.

The comparison-level model also allows the radius to depend on the particular report. The exact local version, Proposition D.1, gives accepted improvement $> \delta_i^u(s,a) - \beta_i^u(s,a)$, termination for nonnegative realized gaps, the same spread bound for a uniform positive gap, and the candidate-specific terminal radius $\delta_i^{\tau(i,a)} + \beta_i^{\tau(i,a)}$, aggregated there into a scan-specific vector-equilibrium guarantee. It does not govern a later resampled report unless a uniform upper bound is declared.

Efficiency of a certified endpoint. Write $C(s) = \sum_i c_i(s)$ for the social cost and $\mathrm{OPT} = \min_{s \in S} C(s)$. Standard $(5/3, 1/3)$-smoothness for unweighted atomic congestion games with nonnegative affine latencies converts a vector-$\boldsymbol{\varepsilon}$ certificate into $C(s) \le (5/2)\,\mathrm{OPT} + (3/2)\sum_i \varepsilon_i$ (Roughgarden, 2015; Christodoulou & Koutsoupias, 2005; Awerbuch et al., 2005). The factor $5/2$ is classical, and once it is fixed the additive coefficient $3/2$ cannot be lowered either: this already follows from the tight multiplicative bound of Christodoulou et al. (2011), though they do not state the additive form. Theorem E.6 exhibits an explicit family attaining this optimal coefficient, and Proposition E.3 shows that dependence on $\sum_i \varepsilon_i$ cannot in general be removed (Theorem E.1 and Corollary E.2). When each worker chooses one tag from a common set, the coverage loss is purely additive.

Proposition 5.2 (Approximate coverage optimality)**.** *In the workboard game, if every worker chooses exactly one tag from a common admissible set, every ε-Nash equilibrium s satisfies $W(s') - W(s) \le (\varepsilon/2)\|\ell(s') - \ell(s)\|_1 \le n\varepsilon$ for every profile s', and both inequalities are tight for every n and ε. Hence every terminal profile of Theorem 4.2 is within $n(\delta+\beta)$ of maximal coverage. Both restrictions are needed; proofs are in Section E.1.*

6 Stale reads and simultaneous rounds

The preceding results treat accepted moves sequentially and absorb stale information into the applicable comparison error. The appendix also derives that error from a sparse interaction structure. For distinct players define the comparison influence

$$\lambda_{ij} := \max(\{0\} \cup \{|\Delta_i(x,a) - \Delta_i(y,a)| : x, y \in S,\ x_{-j} = y_{-j},\ a \in S_i \setminus \{x_i\}\}).$$

If player i reads a profile $r^{i,u}$ whose own coordinate is current, set $D_i^u(s) = \{j \ne i : r_j^{i,u} \ne s_j\}$. When its report has intrinsic error at most $\xi_i^u(s,a)$ at the read profile, Proposition H.1 gives the local certificate

$$|\widehat{\Delta}_i^u(s,a) - \Delta_i(s,a)| \le \xi_i^u(s,a) + \sum_{j \in D_i^u(s)} \lambda_{ij}.$$

Thus stale information enters through a radius specific to each comparison. When a result also imposes an age bound, the read is coordinatewise: different coordinates may come from different recent rounds (Section H). We do not claim the singleton lower bound for a model that requires one coherent historical snapshot.

An exact threshold for one-at-a-time updates. Fix $n \geq 2$, $q \in \{1, \dots, n-1\}$, $L > 0$, and $\eta \geq 0$. Each round activates one player, who scans every alternative at one coordinatewise read of age at most q, with its own coordinate fresh. It chooses a qualifying candidate whenever one exists; every player is activated infinitely often. Suppose scalar scores have pointwise error at most η, and changing one opponent's action changes the activated player's true cost by at most L at every fixed own action. Then Theorem H.7 gives the exact universal threshold for finitely many moves:

$$\delta \geq 2\eta + L \min\{2q, n-1\}.$$

Equality guarantees termination, even with partial scans or unfair schedules; below the threshold, stationary scalar scores can sustain a fair cycle. For $2q \leq n-1$, the proof bounds comparison error and uses potential descent. Beyond that point, Theorem H.5 prevents each player from revisiting an action it left, giving at most $\sum_i (|S_i| - 1)$ accepted moves. The lower construction combines binary and ternary interactions to cover every (n, q), including even n.

This gives a conditional design comparison. With $n = 5$, $\eta = 0.05$, and $L = 0.1$ in common cost units, the threshold is 0.3 at age $q = 1$ and 0.5 at ages $q = 2, 3, 4$. Shortening age three to age two leaves this worst-case margin unchanged; shortening it to one lowers it. A later fresh complete rejected scan certifies radius $\delta + 2\eta$. These follow from specified error and influence bounds; none is an estimate for a deployed judge.

Simultaneous writes need an additional contention charge. In round t, let M_t be the accepted movers. For $i \in M_t$, write δ_i^t, ξ_i^t, and D_i^t for the margin, intrinsic-error radius, and stale-coordinate set of its round-t comparison, and define

$$b_i^t := \delta_i^t - \xi_i^t - \sum_{j \in D_i^t} \lambda_{ij}, \qquad \chi(M_t) := \sum_{\substack{i,j \in M_t \\ i<j}} \lambda_{ij}.$$

Exact potentiality gives $\lambda_{ij} = \lambda_{ji}$: both are maxima of the same absolute mixed potential difference. Each unordered mover pair is charged once. For nonempty rounds, Theorem H.2 proves $\Phi(s^t) - \Phi(s^{t+1}) > \sum_{i \in M_t} b_i^t - \chi(M_t)$. Hence, if the certified slack $\sum_{i \in M_t} b_i^t - \chi(M_t)$ is nonnegative in every nonempty round, finitely many rounds change the state (a count of events, not of time). If the slack is at least $\alpha |M_t|$ with $\alpha > 0$, the total number K of accepted moves satisfies $K < \Delta_\Phi / \alpha$ whenever $K \geq 1$. With fresh reads, at most m movers per round, and η and L as above, the slack is nonnegative whenever $\delta \geq 2\eta + L(m-1)$; for $m \geq 2$ this threshold is exact (Proposition H.3).

Two workers claiming one tag. Take the workboard game of Proposition 3.1 with one tag τ and $S_1 = S_2 = \{\varnothing, \{\tau\}\}$. Let $r_\tau(1) = 3$ and $r_\tau(2) = 1$, so $g_\tau(0) = 0$, $g_\tau(1) = 3$, and $g_\tau(2) = 4$. Use costs $c_i = 3 - u_i$ and cost potential $\Phi = -W$. At $(\varnothing, \varnothing)$, both workers report their exact improvement of 3 and simultaneously claim the tag, with $\delta_i = 2$, $\xi_i = 0$, and $D_i = \varnothing$. Thus $b_1 = b_2 = 2$. Changing the other worker's ownership changes a claiming improvement from 3 to 1, or a dropping improvement from -3 to -1; hence $\lambda_{12} = \lambda_{21} = 2$ and $\chi(\{1, 2\}) = 2$. The actual potential drop is $4 > 2 = b_1 + b_2 - \chi(\{1, 2\})$. Adding the two pre-write improvements instead gives 6, overcounting the drop by 2. This is a mathematical instance of the round bound, not an observation of language-model behavior.

Termination alone does not give an endpoint clock bound. With $R \leq K$ nonempty rounds, read age at most B, and a full scan for each player in every F-round block, Proposition H.8 gives a fresh exhaustive certificate within $(R+1)(B+F)$ rounds. After B quiet rounds, suppose each player still awaiting a post-refresh full scan is offered one in the next round with probability at least $p_{\min} > 0$, given the past. Fix a deterministic cap $\bar{R} \geq R$ in advance. With probability at least $1 - \zeta$, Proposition H.9 gives the horizon $(\bar{R}+1)(B + \lceil p_{\min}^{-1} \log(n(\bar{R}+1)/\zeta) \rceil)$. These fixed-game guarantees exclude changing costs, arbitrary shared overwrites, and candidates that cease to be legal.

7 What realized reports certify

A report exposes only a comparison, and Definition 2.1 is stated in those terms. Fix a player i and an other-player profile s_{-i}, and write $x(a) = (a, s_{-i})$. A comparison of a with b carries a report $\widehat{\Delta}$ against the true difference $\Delta = c_i(x(a)) - c_i(x(b))$, with residual $e = \widehat{\Delta} - \Delta$. An observed residual gives the lower bound $\beta \geq |e|$ (Section M observes $|e| = 5$ for a language model at margin $1/2$);

finitely many small residuals cannot certify a uniform upper bound on β. Recorded signed residuals nevertheless certify the realized trajectory, without predicting unobserved or future comparisons.

Proposition 7.1 (Realized-path certificate)**.** *Call a comparison* recorded *when the run retains its numerical report and the true improvement between the two compared profiles is computable. Consider a closed-loop run with $m \geq 0$ accepted moves, all recorded and indexed by $t = 0, \ldots, m-1$. Let $\widehat{\Delta}_t$ and Δ_t be the reported and true cost improvements for the realized move, and let $e_t = \widehat{\Delta}_t - \Delta_t$. Then*

$$\Phi(s^0) - \Phi(s^m) = \sum_{t=0}^{m-1} \Delta_t \geq \sum_{t=0}^{m-1} (\delta - e_t),$$

with strict inequality when $m \geq 1$. In particular, if $e_t < \delta$ for every accepted move, the realized path strictly decreases the potential at every step even without a certified uniform β.

Suppose a stopping-time scan evaluates every unilateral deviation a, records every comparison, and rejects every reported improvement. With signed residual $e_i(a)$ for that comparison, the terminal profile is an $\varepsilon_{\mathrm{scan}}$-Nash equilibrium for the scanned game, where

$$\varepsilon_{\mathrm{scan}} = \max(\{0\} \cup \{\delta - e_i(a) : i \in N,\ a \in S_i \setminus \{s_i\}\}).$$

This equals zero when every strategy set is a singleton and the scan is empty.

The proof is in Section D.1.

The terminal certificate requires true costs for every residual and a complete scan. It does not apply to a partial scan, and it does not predict how a resampled report would behave. The realized bounds separate a false move from the stronger claim that one global error budget governs every move.

Recorded residuals (a, b, e) admit one shared scalar explanation $d(a) - d(b) = e$ exactly when every closed walk has zero signed sum. This is the classical graph-gradient characterization (Lim, 2020; Jiang et al., 2011), applied in Proposition I.5. Centering each component's range at zero gives the minimum shared pointwise radius: the largest component half-range, or zero for empty data (Proposition I.2). Failure refutes the shared scalar explanation, not the direct model. Success covers only the recorded occurrences.

8 Scope boundary and limitations

The results are conditional. Reports may be unavailable or nonnumeric, need not share one deterministic error bound, and may come from incomplete scans; time variation and statistical dependence are allowed. Recentering requires the scalar-representable subclass; the stationary ceiling also requires one player and a scalar-stationary field, and the residual-graph test checks whether one representative can explain the combined occurrences. The workboard propositions cover only the marginal-credit game. Nothing here establishes exact potentiality in a deployed system or a uniform error bound.

The core model certifies a quiet scan; it gives neither absorption under resampling nor a clock bound. The appendix treats persistent i.i.d. scalar recurrence and its exponential barrier (Section F), and conservative age-free thresholds for simultaneous stale-read rounds (Theorem H.2). Game changes, forgetting, and compression require further specification.

9 Conclusion

Report representation and read protocol determine termination margins. Under β-accurate direct reports, universal termination holds exactly when $\delta \geq \beta$. Fixed scalar scores exclude stationary one-player cycles possible with direct comparisons. For coordinatewise stale scalar evaluations with fresh own coordinates and singleton updates, the exact threshold is $2\eta + L \min\{2q, n-1\}$: potential descent controls short delays, and no-return controls saturation. A fresh complete rejected scan certifies radius $\delta + 2\eta$; a restricted scan certifies only tested alternatives. Residual size at scale, proposal coverage and stale-context effects remain separate empirical questions.

REPRODUCIBILITY STATEMENT

Section A lists the formal results, their assumptions, and their proof locations. Section M reports exact counts from recorded language-model responses. The supplementary archive contains prompts for the primary comparisons and anonymized response records, with offline replay code to reproduce the numerical summaries and figure.

AI USE STATEMENT

Generative AI tools were used for editing and verification; the authors reviewed this material and take full responsibility for the paper.

REFERENCES

Elif Akata, Lion Schulz, Julian Coda-Forno, Seong Joon Oh, Matthias Bethge, and Eric Schulz. Playing repeated games with large language models. *Nature Human Behaviour*, 9:1380–1390, 2025. doi: 10.1038/s41562-025-02172-y.

Carlos Alós-Ferrer and Nick Netzer. The logit-response dynamics. *Games and Economic Behavior*, 68(2):413–427, 2010. doi: 10.1016/j.geb.2009.08.004.

Galit Ashkenazi-Golan, Domenico Mergoni Cecchelli, and Edward Plumb. Simultaneous best-response dynamics in random potential games. arXiv:2505.10378, 2025.

Vincenzo Auletta, Diodato Ferraioli, Francesco Pasquale, Paolo Penna, and Giuseppe Persiano. Convergence to equilibrium of logit dynamics for strategic games. *Algorithmica*, 76(1):110–142, 2016. doi: 10.1007/s00453-015-0025-7.

Baruch Awerbuch, Yossi Azar, and Amir Epstein. The price of routing unsplittable flow. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 57–66. ACM, 2005. doi: 10.1145/1060590.1060599. Journal version: SIAM Journal on Computing 42(1):160–177, 2013, doi 10.1137/070702370.

Baruch Awerbuch, Yossi Azar, Amir Epstein, Vahab S. Mirrokni, and Alexander Skopalik. Fast convergence to nearly optimal solutions in potential games. In *Proceedings of the 9th ACM Conference on Electronic Commerce (EC)*, pp. 264–273, 2008. doi: 10.1145/1386790.1386832.

Maria-Florina Balcan, Avrim Blum, and Yishay Mansour. The price of uncertainty. *ACM Transactions on Economics and Computation*, 1(3):15:1–15:29, 2013. doi: 10.1145/2509413.2509415. URL https://www.cs.cmu.edu/~ninamf/papers/bbm-uncertainty-full.pdf.

Vittorio Bilò, Angelo Fanelli, Laurent Gourvès, Christos Tsoufis, and Cosimo Vinci. Minimizing Rosenthal's potential in monotone congestion games. In *Proc. 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pp. 343–351. IFAAMAS, 2025.

Lawrence E. Blume. The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5(3):387–424, 1993. doi: 10.1006/game.1993.1023.

Ozan Candogan, Ishai Menache, Asuman Ozdaglar, and Pablo A. Parrilo. Flows and decompositions of games: Harmonic and potential games. *Mathematics of Operations Research*, 36(3):474–503, 2011. doi: 10.1287/moor.1110.0500. URL https://doi.org/10.1287/moor.1110.0500.

Ozan Candogan, Asuman Ozdaglar, and Pablo A. Parrilo. Dynamics in near-potential games. *Games and Economic Behavior*, 82:66–90, 2013. doi: 10.1016/j.geb.2013.07.001. URL https://doi.org/10.1016/j.geb.2013.07.001.

Yihang Chen, Yuxiang Chen, Yuxuan Huang, Meng Fang, Weilin Luo, and Jun Wang. Bilevel coordinated reflection: A game-theoretic approach to multi-agent LLM systems. arXiv:2609.02750v1, 2026.

Steve Chien and Alistair Sinclair. Convergence to approximate Nash equilibria in congestion games. *Games and Economic Behavior*, 71(2):315–327, 2011. doi: 10.1016/j.geb.2009.05.004.

George Christodoulou and Elias Koutsoupias. The price of anarchy of finite congestion games. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 67–73. ACM, 2005. doi: 10.1145/1060590.1060600.

George Christodoulou, Elias Koutsoupias, and Paul G. Spirakis. On the performance of approximate equilibria in congestion games. *Algorithmica*, 61(1):116–140, 2011. doi: 10.1007/s00453-010-9449-2. Conference version in ESA 2009, pp. 251–262.

Pierre Coucheney, Stéphane Durand, Bruno Gaujal, and Corinne Touati. General revision protocols in best response algorithms for potential games. In *Network Games, Control and Optimization (NetGCoop)*, 2014. URL https://hal.science/hal-01085077.

Jinhao Duan, Renming Zhang, James Diffenderfer, Bhavya Kailkhura, Lichao Sun, Elias Stengel-Eskin, Mohit Bansal, Tianlong Chen, and Kaidi Xu. GTBench: Uncovering the strategic reasoning capabilities of LLMs via game-theoretic evaluations. In *Advances in Neural Information Processing Systems*, volume 37, pp. 28219–28253. Curran Associates, Inc., 2024. doi: 10.52202/079017-0885. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/3191170938b6102e5c203b036b7c16dd-Abstract-Conference.html.

Glenn Ellison. Learning, local interaction, and coordination. *Econometrica*, 61(5):1047–1071, 1993. doi: 10.2307/2951493. URL https://economics.mit.edu/sites/default/files/publications/Learning%2C%20Local%20Interaction%20and%20Coordination.pdf.

Diodato Ferraioli and Paolo Penna. Imperfect best-response mechanisms. *Theory of Computing Systems*, 57(3):681–710, 2015. doi: 10.1007/s00224-014-9597-x. URL https://doi.org/10.1007/s00224-014-9597-x.

Simon Fischer and Berthold Vöcking. Adaptive routing with stale information. In *Proc. 24th ACM Symposium on Principles of Distributed Computing (PODC)*, pp. 276–283, 2005. doi: 10.1145/1073814.1073868.

Ragavendran Gopalakrishnan, Jason R. Marden, and Adam Wierman. Potential games are necessary to ensure pure Nash equilibria in cost sharing games. *Mathematics of Operations Research*, 39(4):1252–1296, 2014. doi: 10.1287/moor.2014.0651.

Xuehang Guo, Xingyao Wang, Yangyi Chen, Sha Li, Chi Han, Manling Li, and Heng Ji. SyncMind: Measuring agent out-of-sync recovery in collaborative software engineering. In *Proc. 42nd International Conference on Machine Learning (ICML)*, volume 267 of *PMLR*, pp. 20993–21066, 2025.

Yuxuan Huang, Yihang Chen, Zhiyuan He, Yuxiang Chen, Ka Yiu Lee, Huichi Zhou, Weilin Luo, Meng Fang, and Jun Wang. Web2BigTable: A bi-level multi-agent LLM system for internet-scale information search and extraction. arXiv:2604.27221, 2026.

Xiaoye Jiang, Lek-Heng Lim, Yuan Yao, and Yinyu Ye. Statistical ranking and combinatorial Hodge theory. *Mathematical Programming*, 127(1):203–244, 2011. doi: 10.1007/s10107-010-0419-x.

Michihiro Kandori, George J. Mailath, and Rafael Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56, 1993. doi: 10.2307/2951777.

Michael Kearns, Mallesh M. Pai, Ryan Rogers, Aaron Roth, and Jonathan Ullman. Robust mediators in large games. arXiv:1512.02698, 2015. Builds on Rogers and Roth, ACM EC 2014.

Yoav Kolumbus, Menahem Levy, and Noam Nisan. Asynchronous proportional response dynamics: Convergence in markets with adversarial scheduling. In *Advances in Neural Information Processing Systems*, volume 36, pp. 25409–25434, 2023. doi: 10.52202/075280-1104.

Jinlong Lei and Uday V. Shanbhag. Asynchronous schemes for stochastic and misspecified potential games and nonconvex optimization. *Operations Research*, 68(6):1742–1766, 2020. doi: 10.1287/opre.2019.1946. URL https://doi.org/10.1287/opre.2019.1946.

David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. American Mathematical Society, second edition, 2017. doi: 10.1090/mbk/107.

Hanyu Li, Dongchen Li, and Xiaotie Deng. Discovering expert-level Nash equilibrium algorithms with large language models. *Nature Communications*, 17:7168, 2026. doi: 10.1038/s41467-026-74003-1.

Lek-Heng Lim. Hodge Laplacians on graphs. *SIAM Review*, 62(3):685–715, 2020. doi: 10.1137/18M1223101.

Anna M. Maddux, Reda Ouhamma, and Maryam Kamgarpour. Finite-time convergence to an ϵ-efficient Nash equilibrium in potential games. arXiv:2405.15497v1, 2024.

Jason R. Marden and Adam Wierman. Distributed welfare games. *Operations Research*, 61(1): 155–168, 2013. doi: 10.1287/opre.1120.1137. URL https://doi.org/10.1287/opre.1120.1137.

Jason R. Marden, H. Peyton Young, Gürdal Arslan, and Jeff S. Shamma. Payoff-based dynamics for multiplayer weakly acyclic games. *SIAM Journal on Control and Optimization*, 48(1):373–396, 2009.

Qiaozhu Mei, Yutong Xie, Walter Yuan, and Matthew O. Jackson. A Turing test of whether AI chatbots are behaviorally similar to humans. *Proceedings of the National Academy of Sciences*, 121(9):e2313925121, 2024. doi: 10.1073/pnas.2313925121.

Yuqiao Meng, Sakshi Sunil Narvekar, Luoxi Tang, Rupali Rajendra Vaje, Yingxue Zhang, Muchao Ye, and Zhaohan Xi. EquiMem: Calibrating shared memory in multi-agent debate via game-theoretic equilibrium. arXiv:2605.09278v1, 2026. URL https://arxiv.org/abs/2605.09278v1.

Dov Monderer and Lloyd S. Shapley. Potential games. *Games and Economic Behavior*, 14(1): 124–143, 1996. doi: 10.1006/game.1996.0044.

Noam Nisan, Michael Schapira, and Aviv Zohar. Asynchronous best-reply dynamics. In *Internet and Network Economics*, volume 5385 of *Lecture Notes in Computer Science*, pp. 531–538. Springer, 2008. doi: 10.1007/978-3-540-92185-1_59.

Pietro S. Oliveto, Tiago Paixão, Jorge Pérez Heredia, Dirk Sudholt, and Barbora Trubenová. How to escape local optima in black box optimisation: When non-elitism outperforms elitism. *Algorithmica*, 80(5):1604–1633, 2018. doi: 10.1007/s00453-017-0369-2.

Paolo Penna. The price of anarchy and stability in general noisy best-response dynamics. *International Journal of Game Theory*, 47:839–855, 2018. doi: 10.1007/s00182-017-0601-y.

Roland Rodriguez. Emergent coordination in multi-agent systems via pressure fields and temporal decay. arXiv:2601.08129v3, 2026. URL https://arxiv.org/abs/2601.08129v3.

Robert W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2:65–67, 1973. doi: 10.1007/BF01737559.

Tim Roughgarden. Intrinsic robustness of the price of anarchy. *Journal of the ACM*, 62(5):32:1–32:42, 2015. doi: 10.1145/2806883. Conference version in STOC 2009, pp. 513–522.

Haoran Sun, Yusen Wu, Yukun Cheng, and Xu Chu. Game theory meets large language models: A systematic survey. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pp. 10669–10677, 2025. doi: 10.24963/ijcai.2025/1184. URL https://www.ijcai.org/proceedings/2025/1184.

Emanuel Tewolde, Xiao Zhang, David Guzman Piedrahita, Vincent Conitzer, and Zhijing Jin. CoopEval: Benchmarking cooperation-sustaining mechanisms and LLM agents in social dilemmas. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. URL https://arxiv.org/abs/2604.15267.

Adrian Vetta. Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions. In *Proc. 43rd IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 416–425, 2002. doi: 10.1109/SFCS.2002.1181966.

Yitao Wu, Si Shen, Rui Yang, Hong Peng, and Bin Hu. Verify, repair, repeat, or stop? robust stopping for noisy verify-repair loops in LLM agents. arXiv:2607.17641v1, 2026. URL https://arxiv.org/abs/2607.17641v1.

Yi Xu, Laura Ruis, Tim Rocktäschel, and Robert Kirk. Investigating non-transitivity in LLM-as-a-judge. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 69583–69612. PMLR, 2025. URL https://proceedings.mlr.press/v267/xu25w.html.

Xie Yi, Zhanke Zhou, Chentao Cao, Qiyu Niu, Tongliang Liu, and Bo Han. From debate to equilibrium: Belief-Driven Multi-Agent LLM reasoning via Bayesian Nash equilibrium. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 72277–72316. PMLR, 2025. URL https://proceedings.mlr.press/v267/yi25c.html.

Quanyan Zhu. Reasoning and behavioral equilibria in LLM-Nash games: From mindsets to actions. arXiv:2507.08208, 2025.

A PROOF MAP

This appendix supplies the omitted proofs and the supporting extensions.

Formal-result inventory. The entries below state the assumptions, statement and proof location, and main qualification for every substantive result.

1. **Workboard potential.** Assumptions: the finite tag–bundle game and displayed marginal-credit utilities. Statement and proof: Proposition 3.1 in the main paper. It does not model arbitrary shared-memory systems.
2. **Robust alignment.** Assumptions: the other players' utilities are fixed and the claim quantifies over every mover-utility realization. Statement: Proposition 3.2 in the main paper; proof in Proposition L.3.
3. **Scalar representation and recentering.** Assumptions: a complete ordered-pair field on one fixed (i, u, s_{-i}) fiber, scalar integrability, and direct error at most β. Statement and proof: Definition 2.2 and Proposition B.3. The conclusion is pointwise error at most $\beta/2$ after a fiberwise shift and does not extend to partial or nonintegrable fields.
4. **Proposal-relative guarantee.** Assumptions: Definition 2.1, $\delta > \beta$, proposal sets containing the incumbent, and a Q-maximal restricted trajectory with terminal scans over the declared sets Q_i. Statement: Proposition 2.3 in the main paper; proof and full-spread witness: Section G. Full Nash stability additionally requires $Q_i = S_i$.
5. **Accepted-move soundness.** Assumptions: a finite exact cost-potential game, direct β-accuracy, and a move accepted above δ. Statement and proof: Lemma 4.1 in the main paper; restatement and proof: Lemma C.1.
6. **Above-threshold convergence and terminal radius.** Assumptions: the preceding model, maximality, and $\delta > \beta$. Statement and proof: Theorem 4.2 in the main paper; expanded proof: Theorem C.2. The spread-over-margin move bound and terminal-radius formulas are prior art.
7. **Equality boundary.** Assumptions: $\delta = \beta > 0$. Statement and proof: Proposition C.3, summarized in Table 1. Termination holds with no positive spread rate, shown there by a staircase witness.
8. **Sub-threshold stationary cycle.** Assumptions: $0 < \delta < \beta$ and the displayed one-player stationary direct field. Statement and proof: Proposition C.4, summarized in the main paper. The proof is an explicit construction.
9. **Exact class-wide threshold.** Assumptions: finite exact-potential games, $\delta > 0$, and all admissible direct fields. Statement: the iff sentence in the main paper's sub-threshold paragraph, restated in Proposition C.4. Proof: Theorem 4.2 and Propositions C.3 and C.4.
10. **Stationary direct leading constant.** Assumptions: $\delta > \beta$ and the single-player direct staircase family. Statement: Proposition L.1; proof: Proposition C.5. Constant one is an unattained asymptotic supremum even under direct stationarity.
11. **Scalar-stationary single-player ceiling.** Assumptions: one player and one scalar-stationary, scalar-representable, β-accurate field. Statement: Proposition L.2; proof and lower family: Proposition C.6. Termination and the move bound hold for every $\delta > 0$; the ceiling needs $\delta > \beta$; no multiplayer ceiling is claimed.

12. **Player-specific parameters.** Assumptions: player-specific parameters with $\delta_i > \beta_i$ for every player and a maximal trajectory. Statement: Theorem 5.1 in the main paper; proof: the per-player constant specialization of Proposition D.1. Equality and sub-threshold qualifications follow immediately afterward.
13. **Local comparison budgets.** Assumptions: report-specific margins and radii, with either nonnegative gaps or a uniform positive gap for the rate. Statement: Proposition D.1; proof: Section D. The accepted-occurrence condition is pathwise, while the all-occurrence condition is an ex-ante sufficient condition. Terminal radii attach to the actual scan occurrences, aggregate into a scan-specific vector, and do not automatically govern later resampling.
14. **Efficiency of certified endpoints.** Assumptions: a (λ, μ)-smooth game with $\mu < 1$ and a vector-equilibrium certificate. Statement and proof: Theorem E.1 and Corollary E.2. The conversion is the standard smoothness argument; the direct-report contribution is only the certified slack supplied by the preceding results.
15. **Vector-slack dependence.** Assumptions: nonnegative coordinate slacks and affine congestion games; constant latencies suffice. Statement and proof: Proposition E.3. The construction shows that dependence on $\sum_i \varepsilon_i$ cannot be removed; it does not establish the affine coefficient $\frac{3}{2}$.
16. **Exact affine additive constant.** Assumptions: atomic affine congestion games with $n \geq 3$ players and additive equilibrium slack $\varepsilon \in (0, 1)$. Statement and proof: Theorem E.6. The multiplicative factor $\frac{5}{2}$ is classical, and that $\frac{3}{2}$ cannot be lowered once $\frac{5}{2}$ is fixed follows from the tight multiplicative bound of Christodoulou et al. (2011); the displayed family attains both constants explicitly, with positive optimum. The construction's load and deviation calculation is Lemma E.5.
17. **Approximate coverage optimality.** Assumptions: the workboard game with every worker choosing exactly one tag from a common admissible set, and nondecreasing concave coverage functions. Statement: Proposition 5.2 in the main paper; proof, tightness (also for the dynamic consequence), the player-specific version, and counterexamples showing that both restrictions and concavity are needed: Section E.1. The argument is the standard unit-exchange bound. The consequence for dynamics requires an exhaustive terminal scan. These singleton strategies are unrelated to the singleton (one-at-a-time) updates of Theorem H.7.
18. **Persistent scalar resampling.** Assumptions: $0 < \delta < 2\eta$; fresh iid bounded pointwise cost errors in each round, with a full-range law for the tie recurrence and the uniform law for the barrier; persistent reevaluation after null rounds; and the scheduler and selection rules stated in the results. Tie recurrence: Theorem F.2; exponential barrier, with an H-independent floor on positive reachable transition probabilities: Theorem F.3; two-player absorption-time ceiling under almost-sure absorption and a fixed positive transition floor: Proposition F.5. These are scalar-representable stochastic subclass results with $\beta = 2\eta$, not claims about arbitrary stochastic direct fields.
19. **Realized-path and exhaustive-scan certificates.** Assumptions: recorded accepted comparisons; the terminal certificate additionally requires a complete scan of recorded rejected comparisons. Statement: Proposition 7.1 in the main paper; proof: Section D.1. The result applies only to the realized records and gives no uniform or future-error guarantee.
20. **Shared-scalar residual certificate.** Assumptions: one scalar error potential shared across the compared actions, represented by a residual multigraph. Statement and proof: Proposition I.5, from Propositions I.1 to I.3. This is classical graph structure applied to the residuals: nonzero cycles refute that observed scalar explanation, while zero cycles certify neither unobserved pairs nor future reports and do not refute or validate the direct theorem.
21. **Local stale-read certificate.** Assumptions: a fixed true profile, a read profile with the mover's own coordinate fresh, comparison or pointwise influence bounds valid over every intermediate profile, and an intrinsic error bound at the read profile. Statement and proof: Proposition H.1. It derives an occurrence-specific comparison radius.
22. **Sharp uniform stale-read families.** Assumptions: stationary pointwise scalar perceptions, a uniform pointwise influence bound, and the displayed fair column schedule with independent movers. Statement and proof: Theorem H.4. The construction proves the coefficient of the local influence sum sharp, and the low-staleness mB dependence sharp at $n = m(2B + 1)$; it is not an arbitrary heterogeneous-matrix lower bound.

23. **Centered no-return bound.** Assumptions: finite action sets, fresh own coordinates, intrinsic direct error at each accepted read, and at every accepted occurrence a margin exceeding that error by at least the centered opponent variation of Theorem H.5; the terminal radius also needs intrinsically accurate scan reports. Statement and proof: Theorem H.5. It needs neither a potential nor a delay bound. The exact scalar saturated witness is Proposition H.6.
24. **Exact scalar singleton interpolation.** Assumptions: finite exact-potential games, fair singleton exhaustive-scan dynamics, fresh own coordinates, coordinatewise bounded read age, uniform pointwise influence, and scalar pointwise error. Statement and proof: Theorem H.7. The lower perception is stationary and the formula covers every player count and every delay between one and $n-1$. Termination holds without fairness, the exhaustive scan or the commit rule.
25. **Simultaneous stale-read rounds.** Assumptions: a finite exact-potential game, the preceding local certificate for every accepted report, simultaneous coordinate writes in one fixed game, and a pairwise comparison-influence charge among the movers. Statement and proof: Theorem H.2; bounded-activation endpoint certificate, whose radius also needs intrinsically accurate scan reports: Proposition H.8.
26. **Exact contention payment for fresh reads.** Assumptions: the rounds of Theorem H.2 with fresh reads, a common margin, intrinsic error at most β, at most m movers per round, and pointwise influence at most L. Statement and proof: Proposition H.3. The threshold $\delta \geq \beta + L(m-1)$ is exact; below it, all workers of a two-tag workboard can switch together forever; for $m \geq 2$, this is an instance of classical concurrent cycling. The scalar-stationary construction needs $m \geq 2$, and with stale reads the combined payment is not claimed sharp.
27. **High-probability endpoint certification.** Assumptions: a finite deterministic nonempty-round cap $\bar{R}$, freshness after B quiet rounds, and a history-conditional per-round full-scan probability floor $p_i > 0$. Statement and proof: Proposition H.9. No independence assumption or wall-clock bound from qualitative fairness alone is claimed, and the certificate's radius also needs intrinsically accurate scan reports.
28. **System scope.** Assumptions: one fixed game, atomic unilateral moves, continuing candidate availability, maximal stopping for the theorem, and an exhaustive terminal scan for record-based certification. Location: Section H. These are deductions from the definitions; game-changing orchestration is outside scope.

B DEFINITIONS

Let G be a finite game with players N, nonempty finite strategy sets S_i, costs $c_i : S \to \mathbb{R}_{\geq 0}$, and exact cost potential Φ, so every unilateral deviation satisfies

$$\Phi(a, s_{-i}) - \Phi(s) = c_i(a, s_{-i}) - c_i(s).$$

Write $\Delta_\Phi = \max_s \Phi(s) - \min_s \Phi(s)$. Finite exact potentials and the resulting finite-improvement property are classical (Monderer & Shapley, 1996); congestion games supply the canonical construction (Rosenthal, 1973).

Definition B.1 (Direct noisy better response)**.** For $a \neq s_i$, let $\Delta_i(s, a) = c_i(s) - c_i(a, s_{-i})$. Fix $\beta \geq 0$ and $\delta > 0$. A direct report field assigns $\widehat{\Delta}_i^u(s, a) \in \mathbb{R}$ at each comparison occurrence u, with

$$\left| \widehat{\Delta}_i^u(s, a) - \Delta_i(s, a) \right| \leq \beta$$

for every i, u, s, a. The deviation is accepted when $\widehat{\Delta}_i^u(s, a) > \delta$. A trajectory records only accepted moves; its tth move is generated at an occurrence u_t. It is maximal if it is infinite or ends with an exhaustive scan rejecting every nonincumbent unilateral candidate, each at its own occurrence $\tau(i, a)$. Direct stationarity means that each ordered-pair report is independent of its occurrence index.

Remark B.2 (Expected surrogate costs). Outcome randomness can be represented without changing the deterministic game, but only conditionally. Fix a time-invariant conditional law $\mathcal{L}(Y_i \mid s)$ and a nonnegative loss ℓ_i, and set

$$c_i(s) = \mathbb{E}[\ell_i(Y_i) \mid s].$$

If these expected costs satisfy the exact-potential identity, the paper's deterministic results apply at the expectation level; they make no claim about an individual realization of Y_i. The expectation alone

does not establish exact potentiality. If $\widehat{\Delta}_i^u$ is estimated from samples, using the theorems additionally requires a deterministic comparison bound or a declared event on which that bound holds for every comparison used by the run; we prove no finite-sample or high-probability guarantee. A changing outcome law need not change its expected costs. If those costs change, this fixed-game interpretation requires further modeling; augmenting the state must still satisfy the finite exact-potential assumptions. Thus this is an interpretation of c_i, not a stochastic-dynamics extension.

Definition B.1 restates Definition 2.1. The restriction $a \neq s_i$ prevents a null action from being counted as a move. Opposite orientations are distinct reports in the general model.

Proposition B.3 (Scalar representation and recentering)**.** *Fix i, u, s_{-i} and suppose the direct field is specified on every ordered pair of distinct actions in the fiber. It has a scalar representation in the sense of Definition 2.2 if and only if its report sum is zero on every directed closed walk. If it is scalar-representable and β-accurate on the complete fiber, it has a representative whose pointwise error is at most $\beta/2$. Conversely, pointwise error at most η implies direct accuracy with $\beta = 2\eta$.*

Proof. Scalar differences telescope around every directed closed walk. Conversely, fix an anchor r. A singleton fiber is vacuous and is represented by the zero scalar. Otherwise, set $\hat{c}(r) = 0$ and $\hat{c}(a) = \widehat{\Delta}(a, r)$ for $a \neq r$. Two-edge walks give antisymmetry. For distinct a, b different from r, the walk $a \to b \to r \to a$ gives

$$\widehat{\Delta}(a, b) = \widehat{\Delta}(a, r) - \widehat{\Delta}(b, r).$$

The same identity when one endpoint is r follows from the definition and antisymmetry. Thus $\hat{c}$ is a scalar representation. If $d(a) = \hat{c}(a) - c(a)$, direct β-accuracy gives $|d(a) - d(b)| \leq \beta$ for every pair, so the diameter of d is at most β. Subtracting its midrange makes $|d(a)| \leq \beta/2$ without changing any report. The converse is the triangle inequality. □

Completeness is necessary. On a partial three-action path, two residual edges of size β may accumulate to a vertex-error diameter 2β. Cycle consistency on observed edges can therefore diagnose their scalar compatibility, but it cannot supply full-fiber recentering without the missing comparisons.

C Convergence proof and exact threshold

Lemma C.1 (Soundness, restated from Lemma 4.1)**.** *An accepted move $s^t \to s^{t+1}$ by player i satisfies*

$$c_i(s^t) - c_i(s^{t+1}) > \delta - \beta, \qquad \Phi(s^t) - \Phi(s^{t+1}) > \delta - \beta.$$

Proof. Let u_t be the occurrence producing the tth accepted move and write $e_t = \widehat{\Delta}_i^{u_t}(s^t, s_i^{t+1}) - \Delta_i(s^t, s_i^{t+1})$. Acceptance and $e_t \leq \beta$ give $\Delta_i = \widehat{\Delta}_i^{u_t} - e_t > \delta - \beta$. Only player i moves, so the exact-potential identity gives the second claim. □

Theorem C.2 (Convergence and terminal error, restated from Theorem 4.2)**.** *If $\delta > \beta$, every maximal trajectory is finite. Its number m of moves satisfies*

$$m \leq \max\left\{0, \left\lceil \frac{\Delta_\Phi}{\delta - \beta} \right\rceil - 1\right\} \leq \frac{\Delta_\Phi}{\delta - \beta},$$

with $m < \Delta_\Phi/(\delta - \beta)$ whenever $m \geq 1$, and its terminal profile is a $(\delta + \beta)$-Nash equilibrium.

Proof. Every finite prefix of $m \geq 1$ moves satisfies, by telescoping Lemma C.1,

$$\Delta_\Phi \geq \Phi(s^0) - \Phi(s^m) = \sum_{t=0}^{m-1} [\Phi(s^t) - \Phi(s^{t+1})] > m(\delta - \beta).$$

The strict inequality uniformly bounds all finite-prefix lengths, so no trajectory is infinite. Integrality gives the first bound; for $m = 0$ it is immediate, and if $\Delta_\Phi = 0$, soundness rules out even one move.

At a terminal profile s^T, let $\tau(i, a)$ be the occurrence at which the exhaustive scan evaluates candidate a. Rejection gives $\widehat{\Delta}_i^{\tau(i,a)}(s^T, a) \leq \delta$, while its residual is at least $-\beta$. Hence $\Delta_i(s^T, a) \leq \delta + \beta$ for every deviation, which is exactly the approximate-equilibrium condition. This argument allows each comparison in a sequential scan to use its own report field. □

Proposition C.3 (Boundary). *If $\delta = \beta > 0$, every maximal trajectory terminates after at most $|S| - 1$ moves at a 2β-Nash equilibrium. No bound on the number of moves in terms of Δ_Φ and β holds: for every $D > 0$ and integer $K \geq 1$, the one-player staircase with actions $\{0, \ldots, K\}$, $c(j) = \Phi(j) = (K - j)D/K$, stationary direct report $D/K + \beta$ on every unit-forward pair and exact reports elsewhere has $\Delta_\Phi = D$ and a maximal trajectory from 0 with $K = |S| - 1$ moves.*

Proof. By Lemma C.1, every accepted move strictly decreases Φ. Hence no profile repeats; finiteness gives at most $|S| - 1$ moves. The terminal-state calculation in Theorem C.2 remains valid and yields radius $\delta + \beta = 2\beta$. Here $\delta - \beta = 0$, so the spread argument has no positive denominator. In the staircase, each unit-forward report exceeds $\delta = \beta$ with residual exactly β, so the steps $0 \to 1 \to \cdots \to K$ are accepted; at K every exact report is negative, so this trajectory is maximal. The excluded corner $\delta = \beta = 0$ is not admitted by Definition B.1, which assumes $\delta > 0$. □

Proposition C.4 (Stationary cycling below the threshold). *For every $\beta > 0$ and $0 < \delta < \beta$, a one-player, two-action finite exact-potential game and stationary direct field have a trajectory that cycles forever. Thus universal termination holds exactly for $\delta \geq \beta$ over the admitted parameter range.*

Proof. Let both true costs be zero, set $\Phi \equiv 0$, and define $\widehat{\Delta}(A, B) = \widehat{\Delta}(B, A) = \beta$. Both residuals equal β and both reports exceed δ, so the path alternates forever. The two report occurrences sum to 2β, proving that the witness is not scalar-representable. The upper direction follows from Theorem C.2 and Proposition C.3; this stationary witness gives the lower direction. □

The scalar subclass has the same class-wide threshold; its scalar-stationary part has a different one-player boundary, while for $0 < \delta < \beta$ a scalar that varies with the occurrence, alternating $\pm\beta/2$ between two zero-cost actions, still cycles one player. For an explicit multiplayer witness, take two players with actions A, B, zero true costs, and the cycle $(A, A) \to (B, A) \to (B, B) \to (A, B) \to (A, A)$. Give player 1 scalar reports $\hat{c}_1(A, A) = \hat{c}_1(B, B) = \beta/2$ and $\hat{c}_1(B, A) = \hat{c}_1(A, B) = -\beta/2$; give player 2 the opposite signs. Every displayed move then reports improvement β and is accepted when $0 < \delta < \beta$, while pointwise error is at most $\beta/2$. By contrast, one scalar-stationary player always terminates, as Proposition C.6 shows.

Proposition C.5 (Stationary direct sharpness). *For every $\delta > \beta \geq 0$ and integer $K \geq 1$, the stationary direct staircase in Proposition L.1 makes K moves and has $\Delta_\Phi = (K + 1)(\delta - \beta)$.*

Proof. Let $a_K = (\delta - \beta)(K + 1)/K$ and $c(j) = \Phi(j) = (K - j)a_K$. Report $a_K + \beta$ on every unit-forward pair and the exact improvement on every other ordered pair. Each report has residual at most β, and every unit step is accepted. The scheduler follows $0 \to 1 \to \cdots \to K$; at K, every candidate has negative exact improvement, so the path is maximal. It has $m = K$ and $\Delta_\Phi = Ka_K = (K+1)(\delta-\beta)$, giving ratio $K/(K+1) \to 1$. The strict upper bound in Theorem C.2 shows that one is a supremum, not an attained ratio. □

Proposition C.6 (Scalar-stationary ceiling and lower family). *In the scalar-stationary one-player subclass, every nonempty trajectory satisfies $m\delta < \Delta_\Phi + \beta$. In all cases,*

$$m \leq \max\left\{0, \left\lceil \frac{\Delta_\Phi + \beta}{\delta} \right\rceil - 1\right\}.$$

If $\delta > \beta$, the normalized asymptotic supremum is $(\delta - \beta)/\delta$.

Proof. Fix a stationary representative $\hat{c}$. Every nonempty finite prefix of accepted scalar differences telescopes, and the full-pair β bound gives

$$m\delta < \hat{c}(s^0) - \hat{c}(s^m) \leq c(s^0) - c(s^m) + \beta \leq \Delta_\Phi + \beta.$$

This uniform prefix bound excludes an infinite trajectory, and integrality gives the displayed all-case bound, including $m = 0$. For sharpness, let $c(j) = (K - j)a_K^{\mathrm{sc}}$ and

$$\hat{c}(j) = c(j) + \frac{\beta}{2} - \frac{\beta j}{K}, \qquad a_K^{\mathrm{sc}} = \left(\delta - \frac{\beta}{K}\right)\left(1 + \frac{1}{K}\right).$$

The pointwise error lies in $[-\beta/2, \beta/2]$, so all direct residuals are within β, and every unit step reports $a_K^{\mathrm{sc}} + \beta/K > \delta$. Thus $m = K$, while $\Delta_\Phi = K a_K^{\mathrm{sc}}$ and the ratio to $\Delta_\Phi/(\delta - \beta)$ tends to $(\delta - \beta)/\delta$. At action K, every backward candidate has larger reported scalar cost and is rejected, so the trajectory is maximal. □

D LOCAL COMPARISON BUDGETS

Proposition D.1 (Local comparison budgets)**.** *Suppose every admissible report occurrence (i, u, s, a) has margin $\delta_i^u(s, a) > 0$, radius $\beta_i^u(s, a) \geq 0$, and satisfies*

$$\left|\widehat{\Delta}_i^u(s, a) - \Delta_i(s, a)\right| \leq \beta_i^u(s, a).$$

Every accepted move has true improvement greater than $\delta_i^u(s, a) - \beta_i^u(s, a)$. Along any trajectory on which this gap is nonnegative at every accepted occurrence, at most $|S| - 1$ moves are accepted; thus every such maximal trajectory terminates. If every accepted gap on the trajectory is at least $\alpha > 0$, its length m satisfies

$$m \leq \max\left\{0, \left\lceil \frac{\Delta_\Phi}{\alpha} \right\rceil - 1\right\}.$$

More generally, if the t-th accepted move is made by player i_t at occurrence u_t, every trajectory with $m \geq 1$ accepted moves satisfies

$$\sum_{t<m} \left(\delta_{i_t}^{u_t} - \beta_{i_t}^{u_t}\right) < \Phi(s^0) - \Phi(s^m) \leq \Delta_\Phi,$$

with the margins and radii evaluated at the accepted occurrences. With per-player constants this gives the first inequality of Theorem 5.1, and the rate bound above with $\alpha = \min_i(\delta_i - \beta_i)$ gives the second. At an exhaustive terminal scan, candidate a evaluated at occurrence $\tau(i, a)$ has true improvement at most $\delta_i^{\tau(i,a)}(s^T, a) + \beta_i^{\tau(i,a)}(s^T, a)$. Consequently, s^T is a vector-$\boldsymbol{\varepsilon}^{\mathrm{scan}}$ equilibrium for

$$\varepsilon_i^{\mathrm{scan}} := \max\Big(\{0\} \cup \Big\{\delta_i^{\tau(i,a)}(s^T, a) + \beta_i^{\tau(i,a)}(s^T, a) : a \in S_i \setminus \{s_i^T\}\Big\}\Big).$$

Proof. For an accepted occurrence, write $e_i^u(s, a) = \widehat{\Delta}_i^u(s, a) - \Delta_i(s, a)$. Then

$$\Delta_i(s, a) = \widehat{\Delta}_i^u(s, a) - e_i^u(s, a) > \delta_i^u(s, a) - \beta_i^u(s, a).$$

Exact potentiality identifies this with the potential drop. A nonnegative gap therefore gives strict descent and forbids repeated profiles. Finiteness implies termination and the $|S| - 1$ bound. Summing the per-move inequality over any $m \geq 1$ accepted moves gives the displayed sum bound. If every accepted gap is at least $\alpha > 0$, summing gives $\Delta_\Phi > m\alpha$ for every nonempty prefix; integrality gives the stated all-case bound. Finally, rejection of candidate a at its own scan occurrence $\tau(i, a)$ and the lower residual bound give

$$\Delta_i(s^T, a) \leq \delta_i^{\tau(i,a)}(s^T, a) + \beta_i^{\tau(i,a)}(s^T, a).$$

Taking the maximum over the nonincumbent candidates gives the stated coordinate; adjoining 0 makes it well-defined when player i has no such candidate. No common scan time or future persistence is used. □

Pathwise versus ex-ante conditions. The nonnegative-gap hypothesis above is pathwise: it need only hold at report occurrences that the trajectory actually accepts. A stronger ex-ante sufficient condition, uniform over all trajectories, is $\delta_i^u(s, a) \geq \beta_i^u(s, a)$ at every admissible report occurrence. Likewise, an ex-ante rate follows from the uniform lower bound $\delta_i^u(s, a) - \beta_i^u(s, a) \geq \alpha > 0$ at every admissible occurrence.

What the occurrence label may encode. The index u may record a changing prompt, model, retrieval result, or tool context when that context changes only the comparison report, declared margin, and certified radius while the true game and legal deviations remain fixed. Changing only which legal candidates are proposed or exposed is covered by Proposition 2.3. If context changes the true costs or legal deviations, the environment must be frozen for the episode or incorporated into an augmented model that is again finite and exact-potential and whose accepted transitions are unilateral deviations; state augmentation alone is not sufficient.

D.1 Realized-path and exhaustive-scan certificates

Proof of Proposition 7.1. Acceptance gives $\widehat{\Delta}_t > \delta$, hence $\Delta_t = \widehat{\Delta}_t - e_t > \delta - e_t$. Exact potentiality identifies each Δ_t with the realized potential drop. Summing gives the identity and the strict bound for $m \geq 1$; for $m = 0$, both sums and the drop are zero. At a rejected comparison, $\widehat{\Delta}_i(a) \leq \delta$, so its true improvement satisfies $\Delta_i(a) = \widehat{\Delta}_i(a) - e_i(a) \leq \delta - e_i(a)$; maximizing over the exhaustive scan together with zero gives the second claim, including an empty scan. □

E Efficiency of certified endpoints

The convergence results certify how much unilateral improvement remains at a terminal state. Smoothness converts that certificate into a social-cost bound. This step depends only on the equilibrium inequalities, not on whether the reports are direct or scalar-representable.

Write $C(s) = \sum_i c_i(s)$ and $\mathrm{OPT} = \min_{s \in S} C(s)$. A cost game is (λ, μ)-*smooth*, with $\lambda > 0$ and $0 \leq \mu < 1$, when

$$\sum_i c_i(s_i^*, s_{-i}) \leq \lambda C(s^*) + \mu C(s) \qquad \text{for every } s, s^* \in S.$$

This is the standard smoothness condition of Roughgarden (2015).

Theorem E.1 (Efficiency of a vector-equilibrium certificate)**.** *If s is a vector-ε equilibrium of a (λ, μ)-smooth game, then*

$$C(s) \leq \frac{\lambda}{1-\mu}\,\mathrm{OPT} + \frac{1}{1-\mu}\sum_i \varepsilon_i.$$

In particular, if every coordinate has the same slack ε, then the additive term is $n\varepsilon/(1-\mu)$.

Proof. Let s^* minimize C. For each player, the vector-equilibrium inequality for the deviation to s_i^* gives $c_i(s) \leq c_i(s_i^*, s_{-i}) + \varepsilon_i$. Summing and applying smoothness,

$$C(s) \leq \sum_i c_i(s_i^*, s_{-i}) + \sum_i \varepsilon_i \leq \lambda\,\mathrm{OPT} + \mu C(s) + \sum_i \varepsilon_i.$$

Move the term $\mu C(s)$ to the left and divide by $1 - \mu$. □

Throughout the affine-congestion results, players are unweighted and atomic, and each resource latency is $\ell_e(x) = a_e x + b_e$ with $a_e, b_e \geq 0$. These games are $(\frac{5}{3}, \frac{1}{3})$-smooth, and their exact pure price of anarchy is $\frac{5}{2}$ (Christodoulou & Koutsoupias, 2005; Awerbuch et al., 2005). Multiplicative approximate equilibria are studied by Christodoulou et al. (2011); their tight bound already implies that the additive coefficient below cannot be improved once the multiplier $\frac{5}{2}$ is fixed. The additive-slack conversion below is just Theorem E.1; what follows it is an explicit family attaining that optimal coefficient.

Corollary E.2 (Affine congestion endpoints)**.** *Every vector-ε equilibrium of an affine congestion game satisfies*

$$C(s) \leq \tfrac{5}{2}\,\mathrm{OPT} + \tfrac{3}{2}\sum_i \varepsilon_i.$$

Consequently, an exhaustive terminal state of the homogeneous direct dynamics with $\delta > \beta$ satisfies

$$C(s^T) \leq \tfrac{5}{2}\,\mathrm{OPT} + \tfrac{3}{2} n(\delta + \beta).$$

For the local model of Proposition D.1, the same conclusion holds with $\sum_i \varepsilon_i^{\mathrm{scan}}$ in place of $n(\delta + \beta)$.

Proof. Substitute $(\lambda, \mu) = (\frac{5}{3}, \frac{1}{3})$ in Theorem E.1. The two end-to-end statements use the terminal certificates from Theorem 4.2 and Proposition D.1. □

The dependence on the sum of the coordinate slacks cannot be removed, even for constant latencies.

Proposition E.3 (Vector-slack dependence). *For every nonnegative vector $(\varepsilon_i)_{i\in N}$, there is an affine congestion game with a vector-$\boldsymbol{\varepsilon}$ equilibrium s such that*

$$C(s) = \mathrm{OPT} + \sum_i \varepsilon_i.$$

Proof. Give player i two private singleton resources, g_i and b_i, with constant latencies 0 and ε_i. At the profile where every player uses b_i, player i can improve by exactly ε_i, so the profile is a vector-$\boldsymbol{\varepsilon}$ equilibrium and has cost $\sum_i \varepsilon_i$. The all-g_i profile has cost zero, which is optimal because costs are nonnegative. □

We next show that the affine upper bound is attained in both constants. The construction uses an equilibrium slack ε; it does not assume any particular report model. Thus a direct-dynamics endpoint may be substituted only after its terminal slack has been certified.

Definition E.4 (Additive amplifier). Fix $n \geq 3$ and $\varepsilon \in (0, 1)$. The affine congestion game $\mathcal{F}_{n,\varepsilon}$ has players indexed by $\mathbb{Z}_n$ and resources $\{r_0, \ldots, r_{n-1}\} \cup \{t_0, \ldots, t_{n-1}\}$, with

$$\ell_{r_j}(x) = x, \qquad \ell_{t_j}(x) = (1-\varepsilon)x.$$

Player i has the two strategies

$$P_i = \{r_i, r_{i+1}, t_i\}, \qquad Q_i = \{r_{i+2}, t_{i+1}\},$$

where subscripts are modulo n. Write $\mathbf{P} = (P_i)_i$ and $\mathbf{Q} = (Q_i)_i$.

Lemma E.5 (Loads in the additive amplifier). *At $\mathbf{P}$, every r-resource has load two and every t-resource has load one. Hence*

$$c_i(\mathbf{P}) = 5 - \varepsilon, \qquad C(\mathbf{P}) = n(5-\varepsilon).$$

The only deviation of player i costs $5 - 2\varepsilon$, so it improves the cost by exactly ε and $\mathbf{P}$ is an ε-equilibrium. At $\mathbf{Q}$ every resource has load one, and therefore $C(\mathbf{Q}) = n(2-\varepsilon)$.

Proof. At $\mathbf{P}$, resource r_j is used by players j and $j-1$, while t_j is used only by player j. This gives the first display. If player i switches to Q_i, the loads on r_{i+2} and t_{i+1} rise from two to three and from one to two. The deviation cost is $3 + 2(1-\varepsilon) = 5 - 2\varepsilon$, so the improvement is ε. At $\mathbf{Q}$, each r_j and each t_j has one user, giving total cost $n(2-\varepsilon)$. □

Theorem E.6 (The additive constant is exactly $\frac{3}{2}$). *For every $n \geq 3$ and $\varepsilon \in (0, 1)$, the ε-equilibrium $\mathbf{P}$ of $\mathcal{F}_{n,\varepsilon}$ satisfies*

$$\mathrm{OPT}(\mathcal{F}_{n,\varepsilon}) = n(2-\varepsilon) > 0, \qquad C(\mathbf{P}) = \tfrac{5}{2}\,\mathrm{OPT} + \tfrac{3}{2}n\varepsilon.$$

Hence the affine bound in Corollary E.2 is attained in both constants on one nondegenerate instance, and no universal coefficient smaller than $\frac{3}{2}$ can replace its additive coefficient.

Proof. By Lemma E.5, $\mathbf{P}$ is an ε-equilibrium with cost $n(5-\varepsilon)$, while $\mathbf{Q}$ shows $\mathrm{OPT} \leq n(2-\varepsilon)$. Applying Corollary E.2 to $\mathbf{P}$ gives

$$n(5-\varepsilon) \leq \tfrac{5}{2}\,\mathrm{OPT} + \tfrac{3}{2}n\varepsilon,$$

and therefore $\mathrm{OPT} \geq n(2-\varepsilon)$. Thus equality holds in the upper bound supplied by $\mathbf{Q}$. Substitution gives

$$\tfrac{5}{2}n(2-\varepsilon) + \tfrac{3}{2}n\varepsilon = n(5-\varepsilon) = C(\mathbf{P}).$$

The optimum is positive because $\varepsilon < 1$. If a smaller universal additive coefficient existed, applying it to this same family would contradict the displayed equality. □

E.1 Coverage loss of approximate workboard equilibria

This subsection proves Proposition 5.2: the inequality (Proposition E.7), its tightness, the necessity of both restrictions, and the consequence for Theorem 4.2. Throughout, every worker chooses exactly one tag from a common admissible set, each g_τ is nondecreasing and concave with $g_\tau(0) = 0$, and $r_\tau(k) = g_\tau(k) - g_\tau(k-1)$ is therefore nonnegative and nonincreasing in k. Write $\ell(s) = (\ell_\tau(s))_{\tau \in \mathcal{T}}$ for the load vector. In congestion-game terms these are singleton strategies, unrelated to the singleton (one-at-a-time) updates of Theorem H.7. The argument is the standard unit-exchange bound for maximizing a separable concave function of the loads, applied to approximate equilibria. At $\varepsilon = 0$ it gives a direct proof that every pure equilibrium is coverage-optimal, which the paragraph before Proposition 3.1 derives from Bilò et al. (2025, Proposition 1).

Proposition E.7 (Approximate coverage optimality, the inequality of Proposition 5.2)**.** *Let s be an ε-Nash equilibrium with $\varepsilon \geq 0$, and let s' be any profile. Write $x = \ell(s)$ and $y = \ell(s')$. Then*

$$W(s') - W(s) \leq \tfrac{\varepsilon}{2}\|y - x\|_1 \leq n\varepsilon.$$

Proof. Let $A = \{a : x_a > y_a\}$ and $B = \{b : y_b > x_b\}$. Since both profiles place n workers,

$$D := \sum_{a \in A}(x_a - y_a) = \sum_{b \in B}(y_b - x_b) = \tfrac{1}{2}\|y - x\|_1 \leq \sum_{a \in A} x_a \leq n.$$

If $D = 0$ the load vectors, and hence the coverages, coincide. Otherwise

$$\begin{aligned} W(s') - W(s) &= \sum_{b \in B} \sum_{k=x_b+1}^{y_b} r_b(k) - \sum_{a \in A} \sum_{k=y_a+1}^{x_a} r_a(k) \\ &\leq \sum_{b \in B}(y_b - x_b)\, r_b(x_b + 1) - \sum_{a \in A}(x_a - y_a)\, r_a(x_a), \end{aligned}$$

because r_b is nonincreasing, so $r_b(k) \leq r_b(x_b + 1)$ for $k > x_b$, and $r_a(k) \geq r_a(x_a)$ for $k \leq x_a$. Regard each $a \in A$ as $x_a - y_a$ surplus units and each $b \in B$ as $y_b - x_b$ deficit units, and pair the D surplus units with the D deficit units by any bijection. A pair (a, b) contributes $r_b(x_b + 1) - r_a(x_a)$. Since $x_a > y_a \geq 0$, some worker is on a at s. Since $y_b > x_b \geq 0$, some worker chooses b at s', so b lies in the common admissible set and a worker on a may move to b. Its utility at s is $r_a(x_a)$, and because $b \neq a$ its utility after the move is $r_b(x_b + 1)$. In utility form the ε-equilibrium condition reads $u_i(\{b\}, s_{-i}) \leq u_i(s) + \varepsilon$, so $r_b(x_b + 1) - r_a(x_a) \leq \varepsilon$. Each pair uses a separate unilateral deviation from s, so one worker may serve several pairs. Summing over the D pairs gives $W(s') - W(s) \leq D\varepsilon$. □

Player-specific radii. If instead s is a vector-$\boldsymbol{\varepsilon}$ equilibrium, assign the $x_a - y_a \leq x_a$ surplus units of each $a \in A$ to distinct workers on a. Each worker then serves at most one pair, and the same argument gives $W(s') - W(s) \leq \sum_i \varepsilon_i$. Every terminal profile of Theorem 5.1 is therefore within $\sum_i(\delta_i + \beta_i)$ of maximal coverage.

Tightness. Take two tags with $g_a(k) = k$ and $g_b(k) = (1 + \varepsilon)k$, give every worker the common admissible set $\{a, b\}$, and let every worker choose a. A worker's gain from moving to b is exactly ε, so this profile is an ε-Nash equilibrium. Moving every worker to b raises coverage by $n\varepsilon$, while $\frac{1}{2}\|y - x\|_1 = n$, so both inequalities hold with equality for every $n \geq 1$ and $\varepsilon \geq 0$.

Both restrictions are needed. Each example below has $\varepsilon = 0$, where the bound would force every equilibrium to be coverage-optimal. *Bundles of size two.* Let $g_\tau(k) = \min\{k, 1\}$ on tags a, b, c, d, and give both workers the common admissible set $\{ab, ac, ad, bc\}$. The profile (ab, ac) is an exact equilibrium with coverage 3, while (ad, bc) has coverage 4. The failure persists when every two-tag bundle is admissible: with $g_a \equiv 0$, $g_b = g_c = \min\{k, 1\}$ and $g_d(k) = k$, the profile (ad, bc) is an exact equilibrium with coverage 3, while (bd, cd) has coverage 4. *Worker-specific admissible sets.* Let worker 1 choose from $\{a, b\}$ and worker 2 from $\{b, c\}$, with $g_\tau(k) = v_\tau \min\{k, 1\}$ and $(v_a, v_b, v_c) = (2, 3, 1)$. The profile (b, c) is a strict equilibrium with coverage 4, while (a, b) has coverage 5. Concavity, part of the workboard model, cannot be dropped either: with two workers, common admissible set $\{a, b\}$, $g_a(k) = k$ and $(g_b(0), g_b(1), g_b(2)) = (0, 0, 5)$, the profile (a, a) is a strict equilibrium with coverage 2, while (b, b) has coverage 5.

Consequence for dynamics and scope. Pass to the cost form $c_i' = K_i - u_i$, an exact potential game with $\Phi_{\mathrm{wb}} = -W$. Reports and acceptance concern true improvements $u_i(a, s_{-i}) - u_i(s)$, which do not depend on K_i. A maximal trajectory of Theorem 4.2 ends only after an exhaustive scan, so its terminal profile is a $(\delta+\beta)$-Nash equilibrium over all unilateral deviations, and Proposition E.7 with $\varepsilon = \delta + \beta$ gives the bound $n(\delta+\beta)$ of Proposition 5.2. This bound is attained: in the tight family with $\varepsilon = \delta + \beta$, the stationary β-accurate reports $\widehat{\Delta} = \Delta - \beta$ equal δ for every candidate at the profile where all workers choose a, so every candidate is rejected and the zero-move maximal trajectory stops $n(\delta+\beta)$ below maximal coverage. A restricted scan certifies only the alternatives it tests (Proposition 2.3) and need not supply the global ε used here. If the terminal scan reads stale coordinates, its radius must include the staleness term $\sum_{j \in D_i} \lambda_{ij}$ of Proposition H.1; the fresh exhaustive certificates of Propositions H.8 and H.9 need none.

F PERSISTENT I.I.D. SCALAR RESAMPLING

The direct model allows adversarial and time-varying reports, but it does not assign them a probability law. This section studies one narrower stochastic model. It is stated in pointwise scalar costs, so every realized comparison is scalar-representable and has direct error at most $\beta = 2\eta$.

Definition F.1 (Persistent scalar-resampling dynamics)**.** Fix $\eta > 0$, $\delta > 0$, and a distribution ν supported on $[-\eta, \eta]$. In round t, an activated player i draws independent errors $\xi_a^t \sim \nu$ for the incumbent and every candidate $a \in S_i \setminus \{s_i^t\}$, and evaluates the scalar costs

$$c_i(a, s_{-i}^t) + \xi_a^t.$$

A candidate is accepted when its noisy improvement is greater than δ. If several candidates qualify, one is selected according to the declared selection rule. If none qualifies, the round is null and the process continues with fresh errors. The process *settles* when it makes only finitely many state-changing moves. Errors are independent across actions, players, and rounds. A law is *full-range at level* η when its essential infimum and supremum are $-\eta$ and η.

Continuing after a null round matters. If the process stopped at its first quiet draw, a tied two-action state would stop after a geometric number of rounds for every threshold, which would say nothing about recurrence under repeated evaluation.

Theorem F.2 (I.i.d. recurrence at a tie)**.** *Fix $0 < \delta < 2\eta$ and a full-range law ν. Consider the one-player two-action game with both costs equal to zero, and activate the player in every round of Definition F.1. If*

$$p_\nu(\delta) = \Pr(\xi - \xi' > \delta), \qquad \xi, \xi' \overset{\mathrm{iid}}{\sim} \nu,$$

then $p_\nu(\delta) > 0$, the move indicators are iid $\mathrm{Bernoulli}(p_\nu(\delta))$, and almost surely the process makes infinitely many moves, all with zero true improvement. For $\nu = \mathrm{Unif}[-\eta, \eta]$,

$$p_\nu(\delta) = \frac{(2\eta - \delta)^2}{8\eta^2}.$$

Proof. At either action the unique candidate is tied with the incumbent. Its noisy improvement is therefore the difference of two fresh iid errors, so the move probability is $p_\nu(\delta)$ at every round, independently of the state and history. Full range and $\delta < 2\eta$ make this probability positive. The move indicators are iid, so the second Borel–Cantelli lemma gives infinitely many moves almost surely. Each move switches to the other zero-cost action and leaves the constant exact potential unchanged. Under the uniform law, $\xi - \xi'$ has density $(2\eta - |z|)/(4\eta^2)$ on $[-2\eta, 2\eta]$; integrating from δ to 2η gives the formula. □

The next construction shows that even when absorption occurs almost surely, exact potentiality and an H-independent lower bound on positive reachable transition probabilities do not imply polynomial expected time.

Theorem F.3 (An exponential exact-potential barrier)**.** *Fix $\eta > 0$ and $0 < \delta < 2\eta$. There are constants $0 < a < b$, depending only on (η, δ), such that for every integer $H \geq 2$ there is a two-player nonnegative identical-interest exact-potential game with $O(H^2)$ profiles and an initial profile for which the persistent dynamics with iid uniform errors, a uniform scheduler, and uniform selection among accepted candidates has the following properties:*

1. *exactly $H+1$ profiles are reachable;*
2. *each reachable profile has at most two positive-probability state-changing successors, and every positive reachable transition has probability at least a;*
3. *absorption occurs almost surely, and the absorption time T satisfies*

$$\frac{1}{a}\left(\frac{b}{a}\right)^{H-2} \le \mathbb{E}[T] < \left(\frac{a}{(b-a)^2} + \frac{2a}{b-a}\right)\left(\frac{b}{a}\right)^{H}.$$

For $(\eta,\delta)=(1,1)$ one may take $a=1/64$ and $b=9/64$, and then $\mathbb{E}[T]=(5\cdot 9^H-32H-5)/4$ exactly.

Proof. Choose

$$0 < d < \min\{\delta, 2\eta-\delta\}.$$

For $D=\xi-\xi'$ with independent uniform errors, set

$$q_- := \Pr(D>\delta+d), \qquad q_+ := \Pr(D>\delta-d), \qquad a := q_-/2, \qquad b := q_+/2.$$

The triangular tail of D and the choice of d give $0<a<b<1/4$.

Let $r=\lceil H/2\rceil$ and give each player action set $\{0,1,\dots,r\}$. Define the staircase profiles

$$z_{2j}=(j,j), \qquad z_{2j+1}=(j+1,j)$$

whenever the index is at most H. Choose $A>\delta+2\eta$ independently of H, put

$$M_H := A+(H-1)d+(2\eta-\delta)+1,$$

and define a common cost Ψ_H on the full product space by

$$\Psi_H(z_0)=0, \qquad \Psi_H(z_k)=A+(H-k)d \quad (1\le k\le H),$$

with $\Psi_H(s)=M_H$ off the staircase. Give both players the cost $c_1=c_2=\Psi_H$. Every unilateral cost change then equals the change in Ψ_H, so this is a nonnegative identical-interest exact-potential game.

Consecutive staircase profiles differ in one coordinate, and no other pair of staircase profiles are unilateral neighbors. An off-staircase candidate has noisy improvement at most

$$A+(H-1)d-M_H+2\eta<\delta,$$

so it is never accepted. Profile z_0 is absorbing. At each interior staircase profile, the backward and forward neighbors change different players' actions, so each activated player has at most one staircase candidate. For $2\le k\le H-1$, moving from z_k to z_{k-1} raises the common cost by d and moving to z_{k+1} lowers it by d. The uniform scheduler therefore gives these two state-changing transition probabilities as a and b, respectively. At z_H the only state-changing transition has probability a. At z_1, the move to z_0 improves the true cost by $A+(H-1)d>\delta+2\eta$, and hence is accepted whenever its controlling player is activated, with probability $1/2$. Thus precisely the staircase is reachable, absorption occurs almost surely, and all positive transitions, including self-loops, have probability at least a.

Let e_k be the expected time to hit z_1 from z_k. Then $e_1=0$, and first-step analysis gives

$$e_k = 1+ae_{k-1}+be_{k+1}+(1-a-b)e_k \quad (2\le k\le H-1),$$

while $e_H=1+ae_{H-1}+(1-a)e_H$. Put $h_k=e_k-e_{k-1}$. The recurrences become

$$ah_k-bh_{k+1}=1, \qquad h_H=1/a.$$

Backward substitution yields

$$h_k=\sum_{j=0}^{H-k}\frac{b^j}{a^{j+1}},$$

and therefore

$$\mathbb{E}_{z_H}[T]\ge e_H=\sum_{k=2}^{H}h_k\ge\frac{b^{H-2}}{a^{H-1}}=\frac{1}{a}\left(\frac{b}{a}\right)^{H-2}.$$

For the upper bound, put $\rho = b/a > 1$. Summing the geometric series gives $h_k = (\rho^{H-k+1} - 1)/(b-a)$ and $e_H = (a\rho^H - b)/(b-a)^2 - (H-1)/(b-a)$. At z_1 the move to z_0 has probability $1/2$, and the move to z_2, which lowers the common cost by d, has probability b. Every path from z_2 to z_0 passes through z_1, so first-step analysis gives $\mathbb{E}_{z_1}[T] = 1 + b\,(e_2 + \mathbb{E}_{z_1}[T]) + (\frac{1}{2} - b)\,\mathbb{E}_{z_1}[T]$, that is, $\mathbb{E}_{z_1}[T] = 2(1 + bh_2) = 2 + 2(a\rho^H - b)/(b-a)$. The strong Markov property at the first visit to z_1 then gives the exact value

$$\mathbb{E}_{z_H}[T] = e_H + 2(1 + bh_2) = 2 + \left(\frac{a}{(b-a)^2} + \frac{2a}{b-a}\right)\rho^H - \frac{b}{(b-a)^2} - \frac{H - 1 + 2b}{b-a},$$

and the upper bound follows because $b/(b-a)^2 > 1/b > 2$. For $a = 1/64$ and $b = 9/64$ the exact value is $(5 \cdot 9^H - 32H - 5)/4$. The full game has $(\lceil H/2 \rceil + 1)^2 = O(H^2)$ profiles. For $(\eta, \delta, d) = (1, 1, 1/2)$, the triangular-tail formula gives $a = 1/64$ and $b = 9/64$. □

Remark F.4 (Tightness and context). The chain in the proof is a biased birth–death chain, and the exact value is the classical expected-duration computation for such chains (Levin & Peres, 2017, Section 2.5), with the indices reversed and a separate step at z_1. The two bounds of Theorem F.3 therefore differ by the factor $(1 + 2(b-a))/(1 - a/b)^2$, which does not depend on H; for $(a, b) = (1/64, 9/64)$ it is $405/256$. Long waiting times in evolutionary coordination models and exponential mixing times in logit dynamics have classical precedents (Kandori et al., 1993; Ellison, 1993; Auletta et al., 2016), and non-elitist local search bounds hitting times with the same gambler's-ruin method (Oliveto et al., 2018). Our model differs in that bounded errors and a strict margin make the minimizer z_0 absorbing, so the waiting time is an absorption time. The barrier still grows with the instance: with the choice of M_H in the proof, the spread of Ψ_H is linear in H, so relative to it the noise shrinks like $1/H$.

With two players, almost-sure absorption and a fixed positive transition floor imply $\mathbb{E}[T] \le \exp(O(\sqrt{|S|}))$ under this protocol.

Proposition F.5 (A two-player ceiling)**.** *Run the persistent dynamics of Theorem F.3 (uniform scheduler, uniform selection among accepted candidates, iid* $\mathrm{Unif}[-\eta, \eta]$ *errors and* $0 < \delta < 2\eta$*) in a finite game with* n *players, from a profile where absorption is almost sure. Suppose that* $a > 0$*, that every positive transition out of a reachable profile has probability at least* a*, and that* R *profiles are reachable. Then:*

1. *for each player* i *and profile* x*, the set* $\{(y, x_{-i}) : y \in S_i\}$ *contains at most* $1 + 1/(na)$ *reachable profiles;*
2. $\mathbb{E}[T] \le (R-1)a^{-(R-1)}$.

Consequently, with two players, $|S|$ *profiles and a nonabsorbing initial profile,* $\ln \mathbb{E}[T] \le (1 + \frac{1}{2a})\sqrt{|S|}\,\ln\frac{1}{a} + \ln\big((1 + \frac{1}{2a})\sqrt{|S|}\big)$*. The family of Theorem F.3, with* $|S| = (\lceil H/2 \rceil + 1)^2$*, attains this square-root order of* $\ln \mathbb{E}[T]$ *at its fixed parameters, including* $a = 1/64$*; the constants in the exponent need not match.*

Proof. For part 1, let U be the set of reachable profiles in $\{(y, x_{-i}) : y \in S_i\}$. If U is empty there is nothing to prove; otherwise let $x^* \in U$ maximize c_i on U. For $u \in U \setminus \{x^*\}$, the move from x^* to u is a unilateral deviation of player i with true improvement $c_i(x^*) - c_i(u) \ge 0$, so its noisy improvement exceeds δ with probability at least $\Pr(\xi - \xi' > \delta) > 0$, and it is then selected with positive probability. The transition $x^* \to u$ therefore has probability at least a. All these transitions require player i to be activated, which has probability $1/n$, so $(|U| - 1)a \le 1/n$.

For part 2, if $R = 1$ the initial profile is absorbing and $T = 0$. Otherwise, almost-sure absorption forces an absorbing profile to be reachable from every reachable profile, along a path of at most $R - 1$ positive transitions, whose probability is at least a^{R-1}. By the Markov property, $\Pr(T > k(R-1)) \le (1 - a^{R-1})^k$, so $\mathbb{E}[T] = \sum_{t \ge 0} \Pr(T > t) \le (R-1)a^{-(R-1)}$.

With two players and action sets of sizes m_1 and m_2, part 1 applied to the m_2 sets of player-1 deviations gives $R \le m_2(1 + 1/(2a))$, and symmetrically $R \le m_1(1 + 1/(2a))$; also $\min\{m_1, m_2\} \le \sqrt{|S|}$. The upper bound on $\ln \mathbb{E}[T]$ follows from part 2. For the family of Theorem F.3, $\ln \mathbb{E}[T] \ge (H-2)\ln(b/a)$ and $H \ge 2\sqrt{|S|} - 3$, so $\ln \mathbb{E}[T]$ grows at least linearly in $\sqrt{|S|}$. □

These results concern repeated scalar resampling, not the adversarial stopping rule of the core theorem and not unrestricted stochastic direct reports. They show that iid noise does not remove the sub-threshold obstruction and that, for $0 < \delta < 2\eta$, exact potentiality with an H-independent floor on positive reachable transition probabilities does not by itself give expected absorption time polynomial in the number of profiles in this subclass.

G RESTRICTED PROPOSALS

Proof of Proposition 2.3. An accepted move is a genuine unilateral deviation whatever generated it, so Lemma C.1 applies verbatim and each accepted move decreases Φ by more than $\delta - \beta$. Restricting the option set can only remove candidates, never create one, so the telescoping argument of Theorem C.2 bounds the number of moves exactly as before and no restricted trajectory can be infinite. Hence every Q-maximal restricted trajectory is finite. At s^T every nonincumbent $a \in Q_i$ was scanned and rejected. At its own scan occurrence $\tau(i,a)$, $\widehat{\Delta}_i^{\tau(i,a)}(s^T, a) \le \delta$, and the residual bound in Definition B.1 bounds the true improvement by $\delta + \beta$. The guarantee for the incumbent is trivial. If $Q_i = S_i$ for every i this is the Nash condition.

For the converse, one player with actions $\{0,1,2\}$, costs $c(0) = c(1) = M$ and $c(2) = 0$, and $\Phi = c$ is an exact potential with $\Delta_\Phi = M$. Set $Q = \{0,1\}$ and let every step expose $P^u(s) = Q \cup \{s\}$. At incumbent 0, the only proposed and scanned deviation is 1, whose true and reported improvements are both $0 < \delta$, so it is rejected and 0 is terminal for this process. The unscanned action 2 improves the true cost by $M = \Delta_\Phi$, which is arbitrary. Hence no bound better than Δ_Φ holds once the scan is not exhaustive. □

Proposal mechanisms and persistent coverage. The proposition is pathwise: $P_i^u(s)$ may be produced deterministically, adaptively, or by a proposal distribution, and no distributional assumption is needed for descent. We reserve Q_i for the declared terminal scan set; a stochastic implementation may instead denote its proposal law by $q_i^u(a \mid s)$. Persistent coverage recovers the full Nash conclusion only when it is state-conditional. After the last accepted move, the game and profile s^T must remain fixed, proposal and evaluation must continue, and every $a \in S_i \setminus \{s_i^T\}$ must eventually be evaluated at that same profile. The first such occurrences then form a finite exhaustive endpoint scan because each S_i is finite. Coverage at earlier profiles, or full support followed by early stopping, does not suffice. For example, if every player receives infinitely many proposal opportunities at s^T and each untested candidate has history-conditional probability $q_i^u(a \mid s^T) \ge \rho > 0$ at each such opportunity, then coverage holds almost surely. The activation, continued-sampling, and fixed-state clauses are essential.

The two conclusions behave differently. Termination survives any proposal restriction because it rests on a descent argument about accepted moves. Stability does not, because it quantifies over deviations that the process never examined. An observed trajectory can therefore settle while certifying nothing about equilibrium unless its endpoint claim also states the scanned set.

H STALE READS, SIMULTANEOUS UPDATES, AND CERTIFICATION

We first explain when different update frequencies and several local revisions fit the sequential model, then treat stale reads and simultaneous writes.

A slower orchestrator. Definition B.1 places no fairness, rate, or frequency condition on the scheduler, and Theorem C.2 quantifies over *every* maximal trajectory. Thus maximality constrains stopping, not *interleaving*. A component may act arbitrarily rarely, in long bursts, or at very different rates; its moves are a subsequence of the accepted moves, whose order the descent argument does not use. A slower timescale therefore needs no new convergence argument if its updates are unilateral deviations in the same fixed game.

A finite run is certified as maximal only by the exhaustive endpoint scan in Definition B.1; earlier steps face no such requirement. A candidate may be skipped through any number of intervening accepted moves, and if those moves remove its improvement, the terminal scan can reject it without that player having acted earlier.

This is not a liveness or wall-clock statement. The trajectory records accepted moves, so an orchestrator that idles forever without performing the terminal scan leaves only an uncertified finite prefix. Termination also does not require $\delta > \beta$: above the threshold each move drops Φ by more than the positive amount $\delta - \beta$, while at $\delta = \beta > 0$ each move still drops Φ strictly. By Proposition C.3 no profile then repeats, and finiteness of the game implies termination. The guarantee concerns accepted-move counts along maximal runs, not fairness, liveness, or elapsed time.

Atomicity remains an assumption: a write that overwrites a concurrent one is not a unilateral deviation. If the slow component instead changes the game by reassigning tasks, editing the objective, or changing admissible bundles, its update is not a deviation and need not preserve Φ, and the bounds above, which assume a fixed game, do not cover it. A two-timescale extension is therefore needed for an orchestrator that changes the game rather than one that is simply slow.

Agents that do not observe each other's writes. We distinguish four questions: (1) whether a single global state exists; (2) what each agent observes of it; (3) when those observations refresh; and (4) how perceived costs are formed from what was observed.

The dynamics assumes (1), and constrains (4) through the comparison-residual bound. It imposes no condition on (2) or (3). No agent is required to observe the profile, and the direct report field may change arbitrarily between comparison occurrences, so a stale or absent read is not by itself outside the model.

Under three conditions, non-observation enters through (4) and only through (4): each committed update is an atomic unilateral move in the current fixed game; the candidate actions a worker compares stay available to it; and terminality is the exhaustive scan of Definition B.1. All three can fail in a deployed system, with different consequences. A worker that has not read the board may propose a different *set* of alternatives rather than misjudge a fixed pair, which is Proposition 2.3 and not this paragraph: termination and the move bound survive, but stability holds only relative to what was scanned, and the unscanned gap can be as large as Δ_Φ. A run that halts while an *accepted* deviation still exists is not maximal, and the theorem says nothing about it. Under our evidence-based definition, a run that does not scan every deviation is not a fully maximal run either. It may instead be terminal relative to a declared partial set Q under Proposition 2.3, in which case stability is only proposal-relative. Absence of a scan is missing evidence, not evidence of an available move, but it precludes the full terminal certificate. What follows is the residual channel, with those three conditions in force. If worker i has not seen what another worker wrote, the comparison it reports between its own two candidates may be wrong by more than β, and that is the single place the guarantee fails. The quantity to measure is therefore not how often an agent is stale, but how far staleness pushes the comparison residual. A system in which agents rarely refresh but whose local comparisons remain accurate to β is covered as it stands; a system with prompt refreshes whose comparisons are nevertheless distorted is not.

Coordinatewise read age. For statements that impose an age bound, let s^t be the true profile at the start of round t, and set $s^t = s^0$ for $t < 0$. A read $r^{i,t}$ has coordinatewise age at most B if, for every coordinate j, there is a $d_j^{i,t} \in \{0, \dots, B\}$ such that

$$r_j^{i,t} = s_j^{\,t-d_j^{i,t}}.$$

Different coordinates may come from different rounds. A fresh own coordinate means $d_i^{i,t} = 0$. If at most μ coordinates change in each round, such a read differs from s^t on at most $\min\{\mu B, n-1\}$ opponent coordinates.

A local influence certificate for stale reads. For distinct players i and j, define the comparison influence

$$\lambda_{ij} := \max(\{0\} \cup \{|\Delta_i(x,a) - \Delta_i(y,a)| : x, y \in S,\ x_{-j} = y_{-j},\ a \in S_i \setminus \{x_i\}\}).$$

Since $j \neq i$, profiles x and y in this maximum have the same incumbent coordinate for player i, so the same candidate is valid at both. Under exact potentiality, $\lambda_{ij} = \lambda_{ji}$. Indeed, fix the remaining coordinates z, actions $a \neq b$ of player i, and actions u, v of player j. The absolute change in the i comparison is

$$|\Phi(a,u,z) - \Phi(b,u,z) - \Phi(a,v,z) + \Phi(b,v,z)|.$$

This is also the absolute change in the j comparison. Maximizing over the same rectangles gives symmetry: choices with $u = v$ contribute zero, and if either player has only one action both maxima are zero. This argument does not assert symmetry of the pointwise cost influences used below.

Proposition H.1 (Local stale-read certificate)**.** *At occurrence (i, u, s, a), let $r^{i,u} \in S$ be the profile read by player i, with fresh own coordinate $r_i^{i,u} = s_i$, and set*

$$D_i^u(s) := \{j \neq i : r_j^{i,u} \neq s_j\}.$$

Suppose $\xi_i^u(s, a) \geq 0$ and the report is intrinsically accurate at the read profile:

$$|\widehat{\Delta}_i^u(s, a) - \Delta_i(r^{i,u}, a)| \leq \xi_i^u(s, a).$$

Then

$$|\widehat{\Delta}_i^u(s, a) - \Delta_i(s, a)| \leq \xi_i^u(s, a) + \sum_{j \in D_i^u(s)} \lambda_{ij}.$$

Thus the right-hand side is a valid occurrence-specific $\beta_i^u(s, a)$ in Proposition D.1.

For comparison with pointwise influence, define

$$\ell_{ij} := \max(\{0\} \cup \{|c_i(x) - c_i(y)| : x_{-j} = y_{-j}\}).$$

Then $\lambda_{ij} \leq 2\ell_{ij}$. In particular, if $\ell_{ij} \leq L$ and $|D_i^u(s)| \leq q$, the displayed comparison radius is at most $\xi_i^u(s, a) + 2Lq$.

Proof. Order $D_i^u(s) = \{j_1, \ldots, j_q\}$. Starting from $x^0 = s$, obtain x^k by replacing coordinates $j_1, \ldots, j_k$ with their values in $r^{i,u}$. Then $x^q = r^{i,u}$, and consecutive profiles differ only in coordinate j_k. The definition of λ_{ij_k} and the triangle inequality give

$$|\Delta_i(r^{i,u}, a) - \Delta_i(s, a)| \leq \sum_{k=1}^{q} \lambda_{ij_k}.$$

Combining this with intrinsic report accuracy proves the residual bound. For profiles x, y differing only in coordinate j,

$$\begin{aligned} |\Delta_i(x, a) - \Delta_i(y, a)| &\leq |c_i(x) - c_i(y)| \\ &\quad + |c_i(a, x_{-i}) - c_i(a, y_{-i})| \leq 2\ell_{ij}, \end{aligned}$$

which proves the pointwise specialization. □

If a separate delay model allows at most μ coordinate changes per delay step, then a read of age at most B has $q \leq \min\{\mu B, n - 1\}$. Under the uniform pointwise bound L, define $\kappa(B) := L \min\{\mu B, n - 1\}$; the proposition gives the structural residual bound $\beta_i^u \leq \xi_i^u + 2\kappa(B)$. If the intrinsic report is formed by differencing two pointwise estimates, each accurate to η, then $\xi_i^u \leq 2\eta$, so Proposition H.1 gives the $2\eta + 2\kappa(B)$ movement-phase correction. This per-comparison certificate is the input to the simultaneous round extension below. A fresh endpoint scan has $D_i^u(s) = \varnothing$ and therefore uses only its intrinsic error.

Simultaneous accepted writes. Concurrent revision can cycle even in potential games (Nisan et al., 2008; Coucheney et al., 2014; Ashkenazi-Golan et al., 2025), so a simultaneous write needs an additional interaction condition. Consider a round t with true profile s^t and accepted mover set M_t. Each $i \in M_t$ chooses $a_i^t \neq s_i^t$ using a read $r^{i,t}$ with $r_i^{i,t} = s_i^t$, an intrinsic error radius ξ_i^t, and a margin δ_i^t, where

$$|\widehat{\Delta}_i^t - \Delta_i(r^{i,t}, a_i^t)| \leq \xi_i^t, \qquad \widehat{\Delta}_i^t > \delta_i^t.$$

Let $D_i^t := \{j \neq i : r_j^{i,t} \neq s_j^t\}$ and

$$b_i^t := \delta_i^t - \xi_i^t - \sum_{j \in D_i^t} \lambda_{ij}.$$

The next profile s^{t+1} applies every a_i^t, $i \in M_t$, simultaneously. Define the contention charge by summing unordered mover pairs:

$$\chi(M_t) := \sum_{\substack{i,j \in M_t \\ i<j}} \lambda_{ij}.$$

Any conceptual ordering charges a later mover for earlier writes that can change its comparison. Since $\lambda_{ij} = \lambda_{ji}$, every such ordering charges each unordered pair once and has the same total $\chi(M_t)$.

Theorem H.2 (Local descent for simultaneous stale-read rounds)**.** *Every nonempty round satisfies*

$$\Phi(s^t) - \Phi(s^{t+1}) > \sum_{i \in M_t} b_i^t - \chi(M_t).$$

If $\sum_{i \in M_t} b_i^t \geq \chi(M_t)$ in every nonempty realized round, then the true profile changes in only finitely many rounds and is eventually constant. Equality is allowed. If, for some $\alpha > 0$, every nonempty round satisfies

$$\sum_{i \in M_t} b_i^t - \chi(M_t) \geq \alpha |M_t|,$$

then the total number K of accepted moves obeys $K < \Delta_\Phi / \alpha$ whenever $K \geq 1$.

Proof. Fix an ordering $\pi = (i_1, \ldots, i_k)$ and conceptually apply the accepted writes in that order, obtaining profiles $s^{t,0} = s^t, s^{t,1}, \ldots, s^{t,k} = s^{t+1}$. Exact potentiality gives

$$\Phi(s^t) - \Phi(s^{t+1}) = \sum_{\ell=1}^{k} \Delta_{i_\ell}(s^{t,\ell-1}, a_{i_\ell}^t).$$

By Proposition H.1 and strict acceptance, $\Delta_i(s^t, a_i^t) > b_i^t$. Before i_ℓ is conceptually applied, its own coordinate is unchanged and only $i_1, \ldots, i_{\ell-1}$ differ from s^t. Telescoping those coordinates and using the definition of λ gives

$$\Delta_{i_\ell}(s^{t,\ell-1}, a_{i_\ell}^t) > b_{i_\ell}^t - \sum_{h<\ell} \lambda_{i_\ell i_h}.$$

Summing charges each unordered mover pair once and proves the first claim. Under nonnegative certified round slack, every nonempty round strictly decreases Φ; finiteness prevents infinitely many such rounds. Under the α premise, summing the drops gives $\Delta_\Phi > \alpha K$. □

A convenient sufficient condition for the rate premise is

$$b_i^t - \sum_{j \in M_t \setminus \{i\}} \lambda_{ij} \geq \alpha \qquad (i \in M_t),$$

since summing these playerwise charges counts each unordered pair twice. Two uniform specializations of Theorem H.2 make these charges explicit. Suppose each report differences two pointwise estimates accurate to η, each pointwise influence satisfies $\ell_{ij} \leq L$, at most m players move in one round, and every read has age at most B. Then $|D_i^t| \leq \min\{mB, n-1\}$. This is the preceding change-rate budget specialized to $\mu = m$, so $\kappa(B) = L \min\{mB, n-1\}$ and

$$b_i^t \geq \delta - 2\eta - 2\kappa(B).$$

If $\lambda_{ij} = 0$ for every ordered pair of distinct movers, then $\chi(M_t) = 0$ and the inclusive threshold $\delta \geq 2\eta + 2\kappa(B)$ gives strict descent and finite termination; for strict inequality it gives, whenever $K \geq 1$,

$$K < \frac{\Delta_\Phi}{\delta - 2\eta - 2\kappa(B)}.$$

This is the comparison-noninterfering case of Theorem H.2. Cost-level noninterference implies this comparison-level condition, but the converse need not hold.

Without a noninterference restriction, if $k = |M_t|$ then $\lambda_{ij} \leq 2L$ gives $\chi(M_t) \leq 2L\binom{k}{2} = Lk(k-1)$. Hence the more conservative threshold

$$\delta \geq 2\eta + 2\kappa(B) + L(m-1)$$

permits arbitrary simultaneous mover sets. With strict inequality, the corresponding accepted-move bound, valid whenever $K \geq 1$, has denominator $\delta - 2\eta - 2\kappa(B) - L(m-1)$. With fresh reads this contention payment is exact, as the next proposition shows; combined with stale reads it is sufficient and not claimed sharp.

Proposition H.3 (Exact contention payment for fresh reads)**.** *Fix $m \geq 1$, $L > 0$, and $\beta \geq 0$. Consider the rounds of Theorem H.2 with fresh reads, so that every $D_i^t = \varnothing$, a common margin $\delta > 0$, radii $\xi_i^t \leq \beta$, at most m accepted movers per round, and pointwise influences $\ell_{ij} \leq L$. If $\delta \geq \beta + L(m-1)$, the true profile changes in only finitely many rounds, and if the inequality is strict, then $K < \Delta_\Phi/(\delta - \beta - L(m-1))$ whenever $K \geq 1$. Conversely, for every $0 < \delta < \beta + L(m-1)$ there are a workboard game with m workers and two tags, in which $\ell_{ij} = L$ and $\lambda_{ij} = 2L$ for all distinct i, j, and a stationary β-accurate direct report field such that, when every worker reads the current profile in every round, all m workers accept a move in every round and the profile changes in every round. For $m \geq 2$ the reports can be taken scalar-stationary with pointwise error $\eta = \beta/2$, so the fresh-read threshold $\delta \geq 2\eta + L(m-1)$ is also exact.*

Proof. Fresh reads give $b_i^t = \delta - \xi_i^t \geq \delta - \beta$. In a nonempty round with $k = |M_t| \leq m$, the bound $\chi(M_t) \leq Lk(k-1)$ above gives

$$\sum_{i \in M_t} b_i^t - \chi(M_t) \geq k\big(\delta - \beta - L(k-1)\big) \geq k\big(\delta - \beta - L(m-1)\big).$$

Theorem H.2 therefore gives both upper claims, with $\alpha = \delta - \beta - L(m-1)$ in the strict case.

For the construction, every worker chooses one tag from $\{\tau, \sigma\}$, and $r_\tau(k) = r_\sigma(k) = L(m-k)$ for $1 \leq k \leq m$. The common coverage function is then nondecreasing and concave with value 0 at 0. In the cost convention $c_i = L(m-1) - u_i$, a worker whose tag has load ℓ pays $L(\ell - 1) \geq 0$. When an opponent switches tags, the load on a worker's own tag moves by exactly one, so $\ell_{ij} = L$. One load rises by one and the other falls by one, so the worker's improvement from switching changes by $2L$ and $\lambda_{ij} = 2L$. If all m workers share a tag, each pays $L(m-1)$ and would pay 0 alone on the other tag, a true improvement of $L(m-1)$. Let every worker report $L(m-1) + \beta$ at the two shared-tag profiles and its true improvement elsewhere. This field is stationary and β-accurate, and $L(m-1) + \beta > \delta$. Start with every worker on τ. In each round all m workers accept, so the profile alternates between all on σ and all on τ.

For $m \geq 2$ put $\eta = \beta/2$ and let $\hat{c}_i = c_i + \eta$ at the two shared-tag profiles, $\hat{c}_i = c_i - \eta$ at the two profiles in which i is alone on its tag, and $\hat{c}_i = c_i$ elsewhere. The first two sets are disjoint because at a shared-tag profile the load on i's tag is $m \geq 2$. Each reported improvement in the cycle is then $L(m-1) + 2\eta > \delta$. □

Every round of this cycle leaves Φ unchanged, since the two tags exchange their loads, while the certified slack $m(\delta - \beta - L(m-1))$ is negative, as Theorem H.2 allows. For $m \geq 2$, the cycle exhibits the concurrent-revision phenomenon discussed before Theorem H.2. The proposition shows that it matches the uniform payment exactly. For $m = 1$ the construction is the one-player cycle that follows Theorem 4.2, and the threshold is the settling boundary $\delta \geq \beta$. For $m = 2$ and $L = 1$ it is, after relabeling one worker's tags, the coordination example at the end of Section H.1. Adding workers whose only admissible bundle is empty, and extending r_τ and r_σ by zero, gives the same cycle for every $n \geq m$.

The numerical age bound is sufficient but not needed for event-time descent. If the mover's own coordinate is current and its candidate remains legal, then always $|D_i^t| \leq n - 1$, regardless of read age. Hence the age-free saturated thresholds

$$\delta \geq 2\eta + 2L(n-1) \quad \text{and} \quad \delta \geq 2\eta + 2L(n-1) + L(m-1)$$

give finite termination for, respectively, noninteracting and arbitrary simultaneous mover sets. These bounds are conservative: unbounded age need not prevent descent when the actual discrepancy budget is controlled, but it does prevent the fresh endpoint guarantee below without a separate refresh assumption.

H.1 SHARP SCALAR SPECIALIZATIONS OF THE STALE-READ BOUND

The local theorem above is stated for arbitrary comparison influences. The next construction shows that its uniform pointwise specialization is attained at $n = m(2B + 1)$. It covers the one-mover ring and the multi-mover grid in one notation.

Fix integers $m \geq 1$ and $B \geq 1$, put $P := 2B + 1$, and let the players be $N := \{0, \dots, m-1\} \times \mathbb{Z}_P$. For $p = (r, c)$ write $\mathrm{col}(p) := c$. The player has actions $\{0, 1, \top_p\}$ and anchor bit $b_p := \mathbf{1}[c \text{ even}]$. Binarize a profile by replacing $\top_p$ with $1 - b_p$, writing the result as $\bar{s}$ and $T(s) := \{p : s_p = \top_p\}$. For distinct columns define

$$d(c, c') := \min\{(c - c') \bmod P, (c' - c) \bmod P\} \in \{1, \dots, B\},$$

and set

$$\rho_\ell(x, y) := \begin{cases} \mathbf{1}[x \neq y], & \ell \text{ odd}, \\ \mathbf{1}[x = y], & \ell \text{ even}. \end{cases}$$

On binary profiles let

$$\Phi_0(u) := \sum_{\substack{\{p,p'\}: \\ \mathrm{col}(p) \neq \mathrm{col}(p')}} \rho_{d(\mathrm{col}(p),\mathrm{col}(p'))}(u_p, u_{p'}),$$

where every unordered pair is counted once, and write

$$\Sigma_p(u) := \sum_{p': \mathrm{col}(p') \neq \mathrm{col}(p)} (-1)^{d(\mathrm{col}(p),\mathrm{col}(p'))} u_{p'}.$$

Define

$$\begin{aligned} c_p^{\mathrm{bin}}(u) &:= (2u_p - 1)\Sigma_p(u) + mB + 2m\mathbf{1}[B \text{ odd}]u_p, \\ c_p(s) &:= c_p^{\mathrm{bin}}(\bar{s}) + 2mB\mathbf{1}[s_p = \top_p], \\ \Phi(s) &:= \Phi_0(\bar{s}) + 2mB|T(s)|. \end{aligned}$$

Theorem H.4 (The uniform stale-read correction is attained at $n = m(2B + 1)$)**.** *For every $m \geq 1$, $B \geq 1$, $L > 0$, $\eta \geq 0$, and*

$$0 < \delta < 2\eta + 2LmB,$$

there is a finite nonnegative-cost exact-potential game on $n = m(2B+1)$ players, a fair simultaneous-round schedule with exactly m movers per round, coordinatewise reads of age at most B, and stationary scalar perceptions of pointwise error at most η. After the finite ramp, each mover's read differs from the current profile on exactly mB coordinates. For every such stale coordinate j, $\ell_{ij} = L$ and $\lambda_{ij} = 2L$. Other players in the mover's own column have $\ell_{ij} = \lambda_{ij} = 0$, so distinct simultaneous movers have zero comparison influence on one another. The trajectory never terminates; in its periodic phase every accepted move has true improvement zero and reported improvement $2\eta + 2LmB$.

Consequently the coefficient of $\sum_{j \in D_i^t} \lambda_{ij}$ in Proposition H.1 is optimal, and neither the factor two nor the mB dependence in the uniform pointwise specialization of Theorem H.2 can be reduced for $n = m(2B + 1)$. For $m = 1$ this is the crossover $2mB = n - 1$.

Proof. It is enough first to take $L = 1$. For a binary profile and player p, define $D_p\Phi_0(u_{-p}) := \Phi_0(1, u_{-p}) - \Phi_0(0, u_{-p})$. Since $\rho_\ell(1, y) - \rho_\ell(0, y) = (-1)^\ell(2y - 1)$,

$$D_p\Phi_0(u_{-p}) = 2\Sigma_p(u) + 2m\mathbf{1}[B \text{ odd}].$$

This equals $c_p^{\mathrm{bin}}(1, u_{-p}) - c_p^{\mathrm{bin}}(0, u_{-p})$; the bootstrap terms have the same property. Thus $c_p - \Phi$ is independent of s_p, and Φ is an exact potential. There are $2m\lfloor B/2 \rfloor$ positive and $2m\lceil B/2 \rceil$ negative coefficients in Σ_p. Thus $-2m\lceil B/2 \rceil \leq \Sigma_p(u) \leq 2m\lfloor B/2 \rfloor$, and substituting $u_p = 0, 1$ gives $c_p^{\mathrm{bin}}(u) \geq m\mathbf{1}[B \text{ odd}] \geq 0$. The bootstrap surcharge is nonnegative. A switch in another column changes c_p by at most one and attains one, whereas a switch in the same column changes it by zero. The corresponding comparison influence is therefore two across distinct columns and zero within one column.

At round t, let all m players in column $c = t \bmod P$ move. Start from the all-bootstrap profile. During the first P rounds, reads are fresh and column c moves from $\top$ to its anchor bit. Afterwards, a mover in column c reads every player in predecessor column $c - \ell$ immediately before that column's move, for $\ell = 1, \dots, B$, and reads its own and the successor columns freshly. It flips its bit to $\mathbf{1}[t \text{ even}]$. Every read has age at most B, the movers in one column have zero mutual influence, and the schedule is fair.

On a ramp move in column c, let w be the binarized opponent profile. Columns before c carry their anchors and columns after c carry the opposite of their anchors. Grouping the m equal terms in each column gives

$$(-1)^{-c} D_p \Phi_0(w) = m\big([\min\{c, B\} - \max\{c - B, 0\}] - [\min\{2B - c, B\} - \max\{B - c, 0\}]\big) = 0;$$

the two brackets are both c when $c \le B$ and both $2B - c$ otherwise. Thus the bootstrap surcharge makes the ramp improvement exactly $2mB$.

After the ramp every column is constant across its m rows. The predecessor and successor at each distance ℓ carry opposite parity encodings, so their contributions cancel in the true derivative: every intended move has true improvement zero. The stale predecessor value is exactly the opposite of its current value. Hence the two distance-ℓ contributions add in the read profile, yielding observed improvement $2m$ per distance and $2mB$ in total. The wave repeats after $2P$ rounds and changes the profile every round.

For each player, put scalar error $+\eta$ on the finitely many observed incumbent profiles and $-\eta$ on their intended-deviation profiles, and zero elsewhere. The two sets are disjoint. For $m \ge 2$, incumbent observations are column-constant while a deviation splits the mover's column. For $m = 1$, separate profiles first by their bootstrap coordinates. The ramp incumbent is the only evaluation with the mover itself at $\top$; the ramp deviation has $2B - c$ later bootstrap coordinates, whereas a first-wave evaluation has $B - c$ such coordinates when $c < B$. Among all-binary evaluations, the two wave phases differ on every opponent coordinate, while an incumbent and its deviation differ on the mover. The only all-binary ramp deviation, for $c = 2B$, differs from one wave incumbent on the mover and from the other on every fresh successor. Hence no incumbent evaluation is also a deviation evaluation. Thus this is a well-defined stationary scalar perception, and each reported improvement is $2mB + 2\eta > \delta$.

Finally multiply every true cost, potential, and noiseless comparison by L. Pointwise influence becomes L, comparison influence becomes $2L$, and every reported improvement becomes $2LmB + 2\eta$. This proves all claims. □

The preceding construction saturates the comparison perturbation paid by the descent proof. When the same player might revisit one of its own actions, a different telescope gives a smaller saturated correction. For $x \in S_{-i}$ write $\operatorname{osc}_x f := \max_x f(x) - \min_x f(x)$ and define the centered opponent variation

$$\Theta_i := \min_{h : S_{-i} \to \mathbb{R}} \max_{a \in S_i} \operatorname{osc}_x \big(c_i(a, x) - h(x)\big).$$

The minimum exists. Fix one opponent profile x_0 and normalize $h(x_0) = 0$. For any finite objective bound, comparison of a fixed action at x and x_0 bounds every $h(x)$. Each sublevel set is closed and bounded in this finite-dimensional normalized space, hence compact. Since $h = 0$ has finite objective, a minimizer exists.

Theorem H.5 (No return under centered opponent variation)**.** *Consider any revision process in a finite game, with fresh own coordinates. At every accepted occurrence u of player i, from action a to action a', suppose the report is intrinsically accurate at the read opponent profile x^u:*

$$\left| \widehat{\Delta}_i^u - \big(c_i(a, x^u) - c_i(a', x^u)\big) \right| \le \xi_i^u, \qquad \widehat{\Delta}_i^u > \delta_i^u,$$

and suppose $\delta_i^u - \xi_i^u \ge \Theta_i$. Then player i never returns to an action it previously left. Consequently the process has at most $\sum_i (|S_i| - 1)$ accepted moves and terminates. This conclusion needs neither an exact potential, a delay bound, nor noninteracting simultaneous movers.

If, after the last accepted move, every player performs a complete scan at the true terminal profile, rejects every deviation, and reports each candidate a at its scan occurrence $\tau_i(a)$ with intrinsic error at most $\xi_i^{\tau_i(a)}$, then that profile is a vector approximate equilibrium with

$$\varepsilon_i = \max\Big(\{0\} \cup \Big\{ \delta_i^{\tau_i(a)} + \xi_i^{\tau_i(a)} : a \in S_i \setminus \{s_i\} \Big\}\Big),$$

where $\tau_i(a)$ is the scan occurrence for candidate a. Thus $\varepsilon_i = 0$ when player i has no alternative action.

Proof. Fix i and a minimizer h in the definition of Θ_i. Suppose the first return in player i's accepted action sequence is $a_0 \to a_1 \to \cdots \to a_r = a_0$. Let x^ℓ be the read opponent profile used for $a_\ell \to a_{\ell+1}$, and put $f_a(x) := c_i(a, x) - h(x)$. Strict acceptance gives

$$g_\ell := c_i(a_\ell, x^\ell) - c_i(a_{\ell+1}, x^\ell) > \delta_i^\ell - \xi_i^\ell \geq \Theta_i.$$

Cyclically reindexing the second sum yields

$$\sum_{\ell=0}^{r-1} g_\ell = \sum_{\ell=0}^{r-1} \big(f_{a_\ell}(x^\ell) - f_{a_\ell}(x^{\ell-1})\big) \leq r\Theta_i,$$

where indices are modulo r. This contradicts the strict lower bound on every g_ℓ. Hence no return occurs, and finiteness of each action set gives the move bound. At a complete terminal scan, rejection says $\widehat{\Delta}_i^{\tau_i(a)} \leq \delta_i^{\tau_i(a)}$; intrinsic accuracy at the true profile gives $\Delta_i(s, a) \leq \delta_i^{\tau_i(a)} + \xi_i^{\tau_i(a)}$, proving the final claim. □

Choosing $h = 0$ shows

$$\Theta_i \leq \max_{a \in S_i} \operatorname{osc}_x c_i(a, x) \leq \sum_{j \neq i} \ell_{ij}.$$

Thus pointwise scalar reports of error at most η_i satisfy the theorem with $\xi_i^u \leq 2\eta_i$ and, under $\ell_{ij} \leq L$, with the homogeneous threshold $\delta_i \geq 2\eta_i + L(n-1)$. This is strictly smaller than the age-free descent threshold $2\eta_i + 2L(n-1)$. It is a scalar specialization only through the intrinsic bound $2\eta_i$; the no-return proof itself uses direct comparison errors.

Proposition H.6 (The centered scalar cap is exact)**.** *For every $n \geq 2$, $L > 0$, $\eta \geq 0$, and $0 < \delta < 2\eta + L(n-1)$, there are a finite nonnegative-cost exact-potential game, a fair singleton schedule, reads of maximum age $n-1$, and stationary scalar perceptions of pointwise error at most η with a nonterminating trajectory. Every recurrent move has true improvement zero and reported improvement $2\eta + L(n-1)$, and every player satisfies $\Theta_i = L(n-1)$. Every ordered pair of distinct players satisfies $\ell_{ij} = L$. Hence the coefficient and the cap in the preceding scalar specialization are both exact; equality is on the terminating side because acceptance is strict.*

Proof. Give each player actions $\{0, 1, 2, \top\}$, with arithmetic on ordinary actions modulo three and encoding $\bar{\top} := 2$. For $i < j$ set $F(x, y) := -L\mathbf{1}[y = x+1 \pmod 3]$. Choose $R > \delta + \eta + L(n-1)$ and define

$$\Phi(s) := \sum_{i<j} F(\bar{s}_i, \bar{s}_j) + R|\{i : s_i = \top\}|,$$

$$c_i(s) := L(n-1) + \sum_{j<i} F(\bar{s}_j, \bar{s}_i) + \sum_{j>i} F(\bar{s}_i, \bar{s}_j) + R\mathbf{1}[s_i = \top].$$

Only terms incident to i and its own bootstrap surcharge depend on s_i, so $c_i - \Phi$ is independent of s_i. The game is therefore an exact potential game, and the offset $L(n-1)$ makes every cost nonnegative. One opponent switch changes one incident term by at most L, and this value is attained.

For the centered variation, $h = 0$ gives $\Theta_i \leq L(n-1)$. For the reverse inequality, compare two opponent profiles: in x, every $j < i$ plays 2 and every $j > i$ plays 1; in y, every $j < i$ plays 0 and every $j > i$ plays 2. Across these profiles, the cost difference for own action 0 is $-L(n-1)$ and for own action 1 is $+L(n-1)$. Subtracting any common $h(x) - h(y)$ cannot make both absolute differences smaller than $L(n-1)$. Thus $\Theta_i = L(n-1)$.

Start at $(\top, \ldots, \top)$. Offer players in order and move each freshly from $\top$ to 0. Thereafter repeat phases $k = 0, 1, 2$: in phase k, offer players in order and move k to $k+1$. At player i's recurrent move, read each earlier player immediately before its current-phase move and each later player immediately before its preceding-phase move. These are genuine past profiles, including in the first recurrent phase, and their ages are at most $n-1$.

At a recurrent move, earlier players are truly at $k+1$ and later players at k. No incident pair term is active either before or after the true move, so the true improvement is zero. In the read profile, earlier

players appear at k and later players at $k-1$. Moving from k to $k+1$ activates all $n-1$ incident terms at value $-L$, so the noiseless observed improvement is $L(n-1)$.

For each player let O_i be its finite set of recurrent incumbent read profiles and D_i the corresponding intended-deviation profiles. Put scalar error $+\eta$ on O_i, $-\eta$ on D_i, and zero elsewhere. This is well defined: an incumbent in phase k has earlier, own, and later values $(k, k, k-1)$, while its deviation has own value $k+1$; equality between an incumbent from phase k and a deviation from phase ℓ would force $k = \ell$ from a nonempty earlier or later side and $k = \ell + 1$ from the own coordinate. The same encoded argument covers the first recurrent phase. During the ramp, the bootstrap surcharge dominates all incident terms and the possible η error, so each intended move is accepted. Every recurrent reported improvement is $L(n-1) + 2\eta > \delta$, so the period-three sweep runs forever. The schedule is fair, singleton, and changes the true profile in every recurrent round. □

For odd n, the singleton case of Theorem H.4 at $B = (n-1)/2$ already attains the saturated value of Proposition H.6. The next theorem gives a family matching its threshold at every delay $q \in \{1, \dots, n-1\}$, including every even n.

Singleton exhaustive-scan convention. In the next result, one player is activated per round and uses one coordinatewise read in the sense defined above. The activation evaluates every nonincumbent action against that read. If at least one report exceeds the margin, an arbitrary declared selection rule chooses one qualifying candidate; otherwise the round is null. The schedule is fair when every player is activated infinitely often; fairness is a liveness requirement on the orchestration protocol, not a behavioral property of the agents. Termination means that only finitely many state-changing moves occur; the scheduler may continue to produce null rounds.

Theorem H.7 (Exact scalar singleton correction at every delay)**.** *Fix $n \geq 2$, $q \in \{1, \dots, n-1\}$, $L > 0$, and $\eta \geq 0$. Consider fair singleton exhaustive-scan dynamics under the preceding coordinatewise-age convention in finite exact-potential games, with fresh own coordinates, read age at most q, pointwise influence at most L, and scalar perceptions of pointwise error at most η. The exact universal termination threshold is*

$$\delta \geq 2\eta + L \min\{2q, n-1\}.$$

At equality every trajectory terminates. Below it, a finite nonnegative-cost exact-potential game and a stationary perception system admit a fair nonterminating trajectory whose maximum actual read age is exactly q. After a finite ramp, every accepted move has true improvement zero and reported improvement $2\eta + L \min\{2q, n-1\}$. The same threshold remains exact for termination if scans may be partial, qualifying candidates may be declined, or the schedule is unfair.

Proof. For the upper bound, during any q consecutive singleton rounds at most q opponent coordinates change. Hence a coordinatewise age-q read differs from the current profile on at most q opponent coordinates. If $2q \leq n-1$, the pointwise specialization of Proposition H.1 gives comparison error at most $2\eta + 2Lq$, and the singleton case of Theorem H.2 gives strict potential descent at and above the displayed threshold. If $2q > n-1$, Theorem H.5 applies because $\Theta_i \leq L(n-1)$. Both arguments include equality because acceptance is strict.

For the lower bound, put $P := n$ and

$$C_q := L \min\{2q, P-1\}, \qquad a := \min\{q, P-1-q\}.$$

Player $i \in \mathbb{Z}_P$ has regular actions $(x_i, z_i) \in \mathbb{Z}_2 \times \mathbb{Z}_3$ and a bootstrap action $\top_i$, encoded as $(1, 2)$. For every $c \in \mathbb{Z}_P$ and $\ell \in \{1, \dots, a\}$, include the unordered binary edge $\{c, c+\ell \pmod P\}$. Let $w_{c,\ell} := \mathbf{1}[c + \ell \geq P]$ and give this edge the term

$$A_{c,\ell}(x_c, x_{c+\ell}) := L\mathbf{1}[x_c \oplus x_{c+\ell} = w_{c,\ell}].$$

If $2q > P-1$, include every remaining unordered pair $\{i, j\}$, $i < j$, as a ternary edge with term

$$F_{ij}(z_i, z_j) := -L\mathbf{1}[z_j = z_i + 1 \pmod 3];$$

otherwise there are no ternary edges. Each player has $2a$ binary neighbors and

$$d_B := \begin{cases} 0, & 2q \leq P-1, \\ P-1-2a, & 2q > P-1, \end{cases}$$

ternary neighbors, so $L(2a + d_B) = C_q$.

Let Φ_0 be the sum of all edge terms. Choose $H > \delta + C_q + \eta$ and define

$$\Phi(s) := \Phi_0(\bar{s}) + H|\{i : s_i = \top_i\}|,$$

$$c_i(s) := Ld_B + \sum_{e \ni i} \Phi_{0,e}(\bar{s}) + H\mathbf{1}[s_i = \top_i].$$

Only incident edge terms and the own bootstrap term depend on s_i, so $c_i - \Phi$ is independent of s_i. Hence this is an exact-potential game. Binary terms lie in $\{0, L\}$, ternary terms in $\{-L, 0\}$, and the offset Ld_B makes every cost nonnegative. One opponent switch changes one incident term by at most L, and the value L is attained.

Start at the all-bootstrap profile. In ramp rounds $c = 0, \ldots, P - 1$, offer player c, read freshly, and move $\top_c$ to $(0, 0)$. In wave sweep $k \geq 0$ and round c, the player moves from $(k \bmod 2, k \bmod 3)$ to $((k+1) \bmod 2, (k+1) \bmod 3)$. For $d = 1, \ldots, q$, read predecessor $c - d \pmod P$ at its pre-move profile from d rounds earlier, and read every other opponent freshly. These are valid histories even in the first sweep: a wrapped predecessor is then read at its ramp pre-move profile, whose encoding is phase -1. The largest actual age is q. Use the fixed selection rule that chooses $(0, 0)$ from a bootstrap action and, from a regular phase, chooses the joint binary–ternary phase successor whenever that candidate qualifies; break any remaining ties arbitrarily.

Fix a wave move of c. For a binary distance $\ell \leq a$, its stale predecessor and fresh successor are both active at the observed incumbent and both become inactive at the observed deviation, giving observed improvement $2L$. In the true profile the predecessor contributes $-L$ and the successor $+L$, so their true contributions cancel. Every ternary neighbor is stale: if its shorter cyclic distance is $\ell > a = P - 1 - q$, then both directed distances are at most q, including the even antipode. Its oriented ternary term contributes observed improvement L and true improvement zero. Thus every wave move has

$$\Delta_i^{\text{true}} = 0, \qquad \Delta_i^{\text{read}} = 2La + Ld_B = C_q.$$

For fixed i, let $O_i(k)$ and $D_i(k)$ be the observed incumbent and intended-deviation profiles in sweep k. Their unions are finite because the binary and ternary phases repeat every six sweeps. They are disjoint. If a bootstrap symbol occurs, both equal profiles would have to come from the first sweep, but their own coordinates are phases 0 and 1. Otherwise, equality on any opponent's ternary coordinate forces the two sweep indices to agree modulo three, while equality on the mover forces them to differ by one modulo three. Define the stationary perception by adding $+\eta$ on O_i, adding $-\eta$ on D_i, and adding zero elsewhere. Every wave report is then $C_q + 2\eta > \delta$. Ramp moves also qualify: removing the surcharge H can lose at most C_q in edge terms and at most η in perception.

The wave changes one action every round and repeats after six sweeps, so it is nonterminating and fair. This proves the lower bound. For completeness, the same instance has $\Theta_i = C_q$: $h = 0$ gives the upper bound, while the two own actions $(0, 0)$ and $(1, 1)$ have a cost difference of opponent-profile oscillation $2C_q$, forcing every common centering to leave one of their oscillations at least C_q. □

What the upper bound uses. Apart from the standing hypotheses on the game and the perceptions, the upper-bound argument uses only that each accepted report exceeds the margin at a coordinatewise read of age at most q with a fresh own coordinate, and that a round changes at most one coordinate. It therefore applies verbatim when an activation evaluates only some of its alternatives, possibly none, or declines a qualifying candidate, and under any schedule that activates one player per round. The lower-bound trajectory satisfies these weaker rules as well, so the threshold remains exact for this larger class of runs. Fairness, the exhaustive scan and the commit rule serve the endpoint instead: together they eventually produce the fresh complete rejected scan that certifies radius $\delta + 2\eta$.

From accepted-move time to a terminal certificate. Qualitative fairness can eventually supply a fresh endpoint scan after the last accepted move, but cannot bound when it happens. For a quantitative statement, assume that reads have age at most B (with initial prehistory fixed at s^0), and every block of F consecutive rounds contains a full-scan offer for every player. A full scan evaluates every unilateral candidate against that player's read, which equals the current profile after B quiet rounds. It immediately commits an accepted candidate if one exists, and otherwise records that player's rejection scan. The protocol halts once every player has recorded a rejection scan during the same

unchanged post-refresh epoch. Bounded-gap liveness assumptions of this kind are also used in other asynchronous models (Kolumbus et al., 2023).

Proposition H.8 (Bounded endpoint-certification time)**.** *Suppose Theorem H.2 bounds the total accepted moves by K, and let $R \le K$ be the number of nonempty rounds. By the end of at most*

$$(R+1)(B+F) \le (K+1)(B+F)$$

rounds, the protocol halts at a profile carrying a fresh exhaustive endpoint certificate. If $\tau(i,a)$ is the fresh rejection occurrence for candidate a and its report is intrinsically accurate within $\xi_i^{\tau(i,a)}$, then

$$\Delta_i(s^\star, a) \le \delta_i^{\tau(i,a)} + \xi_i^{\tau(i,a)}.$$

A player with no alternative action has terminal radius zero.

Proof. Start at initialization or immediately after a nonempty round. If the next $B+F$ rounds contain no accepted move, the profile is constant. After the first B rounds, every admissible read is fresh; during the following F rounds, every player receives a full scan. With no accepted move, every candidate in those fresh scans is rejected, and intrinsic accuracy gives the displayed terminal bound; once every player has such a scan, the stopping rule emits the certificate and halts. Otherwise a nonempty round occurs within the block and the argument restarts immediately afterward. There are at most R such restarts and then one final certificate block. □

A probabilistic liveness condition gives an analogous end-to-end clock bound without independence. It still restricts the scheduler and therefore is not a consequence of qualitative fairness.

Proposition H.9 (High-probability endpoint-certification time)**.** *Suppose a deterministic integer $\bar{R} \ge 0$ bounds the number of nonempty rounds for every realization covered by Theorem H.2. After initialization and after each nonempty round, call the ensuing interval a quiet epoch. Assume that reads become fresh after B quiet rounds. For each player i, once B quiet rounds have elapsed, conditional on the entire past, if the current epoch has not ended and i has not yet received a post-refresh full scan, suppose each next round offers such a scan with probability at least $p_i > 0$. Let $p_{\min} := \min_i p_i$. For every $\zeta \in (0,1)$, set*

$$h_\zeta := \left\lceil \frac{\log(n(\bar{R}+1)/\zeta)}{p_{\min}} \right\rceil.$$

With probability at least $1-\zeta$, by the end of at most

$$(\bar{R}+1)(B+h_\zeta)$$

rounds the protocol halts at a profile carrying a fresh exhaustive endpoint scan. If every terminal scan report satisfies the intrinsic-accuracy condition of Proposition H.8, the endpoint has the radius stated there.

Proof. The cap $\bar{R}$ fixes the horizon and failure allocation before the epochs are observed. Condition on the history at the start of one quiet epoch. For player i, consider the event that this epoch survives its first $B+h$ rounds and i receives no post-refresh scan in those last h rounds. At each surviving, unscanned history, the next no-scan-and-survival probability is at most $1-p_i$. Iterated conditioning therefore bounds this event by $(1-p_i)^h$, without conditioning on future survival. For $h = h_\zeta$, a union bound over players gives at most $ne^{-p_{\min}h_\zeta} \le \zeta/(\bar{R}+1)$ for an epoch that neither ends nor supplies the certificate within $B+h_\zeta$ rounds. There are at most $\bar{R}$ nonempty rounds and hence at most $\bar{R}+1$ adaptively generated epochs. A second union bound proves the claim. □

Under the nonnegative round-slack premise on every covered realization, strict descent gives $\bar{R} = |\Phi(S)| - 1 \le |S| - 1$, since each nonempty round reaches a new potential value. If the fixed potential image lies in $a + \omega\mathbb{Z}$ for $\omega > 0$, each drop is at least ω, giving the further cap $\lfloor \Delta_\Phi/\omega \rfloor$. A uniform positive per-move margin α gives $\max\{0, \lceil \Delta_\Phi/\alpha \rceil - 1\}$, since $R \le K$. One may take the minimum of these deterministic caps. For the binary-coverage workboard, $r_\tau(1) = 1$ and $r_\tau(k) = 0$ for $k \ge 2$, the potential is minus the number of covered tags; hence $\bar{R} = |\mathcal{T}|$ is valid under the same descent premise. Any deterministic accepted-move cap also bounds R. All caps must be fixed before

observing the offers; the realized round or move count cannot replace this preassigned confidence budget.

These are stopping-time certificates, not absorption claims under continued resampling. If the protocol ignored its stopping rule, a later report could accept a candidate that an earlier occurrence rejected.

Both additional premises are needed for the stated guarantees. In the two-player coordination game with common cost and potential zero on matching profiles and one on mismatches, both players strictly improve by switching from $(0, 1)$, but the simultaneous update reaches $(1, 0)$ and produces a period-two cycle. The relevant comparison changes from $+1$ to -1, so the pairwise contention charge is two. Separately, a qualitatively fair scheduler may postpone one player's first offer for an arbitrarily large finite time and then offer it forever; no finite wall-clock bound follows without a quantitative activation condition. Bounded gaps give Proposition H.8; a conditional probability floor gives Proposition H.9. We present the local certificate, the pairwise contention charge and the endpoint-time bounds as a scoped extension, without a priority claim.

Concurrent bilevel work fixes a task decomposition and local worker payoffs, whose sum is an exact potential. Bounded coupling controls the discrepancy between unilateral payoff changes and changes in the chosen system objective, which is analyzed as an approximate potential (Chen et al., 2026). Its deterministic verifier-gated memory update accepts only a strict decrease in a fixed grounded evaluation risk. Here the game has an exact potential and the residual measures error in reported cost improvements. The system-objective discrepancy and the reporting residual are distinct assumptions, although both analyses use potential-improvement arguments.

I What a family of residuals identifies

The general bound $|e| \le \beta$ treats each direct residual on its own. If a single scalar perception is shared across a set of comparisons, however, their residuals are linked rather than arbitrary. That extra structure gives both a pointwise-error floor and a falsification test. Everything in this section concerns scalar compatibility of the *measurement instrument*; none of it strengthens or weakens Theorem C.2.

Fix a player i and a profile s_{-i} of the others, and write $x(a) = (a, s_{-i})$ for $a \in S_i$. A comparison of the ordered pair (a, b) reports $\widehat{\Delta}$ at its own step t, against the truth $\Delta = c_i(x(a)) - c_i(x(b))$, so its residual is generally just

$$e = \widehat{\Delta} - \Delta.$$

The whole section asks what follows under the extra hypothesis that one scalar report $\hat{c}$ generated the set. Then $e = d(a) - d(b)$ for $d(a) = \hat{c}(x(a)) - c_i(x(a))$. Definition B.1 does not assume either scalar representability or sharing across occurrences, and Propositions I.2 and I.3 are conditional on both throughout.

Record the observations as a finite *multiset* E of oriented edge occurrences $q = (a, b, e)$, one per response, on vertex set $V \subseteq S_i$. Repeating a comparison or reversing it adds a new occurrence rather than overwriting one. Let H be the underlying undirected multigraph: each occurrence joins a and b but retains the stored orientation $a \to b$ and value e. A walk may traverse an occurrence in either direction. Its signed residual contribution is $+e$ in the stored direction and $-e$ in the reverse direction. Thus two parallel occurrences form a two-edge closed walk, whether their stored orientations agree or oppose one another; a self-loop is a one-edge closed walk. Say $d : V \to \mathbb{R}$ *explains* the observations when $d(a) - d(b) = e$ for every occurrence $(a, b, e) \in E$.

Relation to graph Hodge theory. The standard graph characterization says that edge values are differences of a vertex function exactly when every closed walk has zero signed sum. In the language of graph cochains, such an edge flow is exact (a gradient flow); see, for example, the treatment of graph gradients and Hodge decompositions in Lim (2020). Proposition I.1 applies this characterization to residuals. A directly relevant precedent is HodgeRank, which maps paired comparisons to edge flows and separates a globally consistent gradient component from components representing inconsistency (Jiang et al., 2011).

Two things differ, and neither is a claim of novel mathematics. We do not use the least-squares projection studied in HodgeRank. Conditional on exact explainability, we instead minimize the supremum norm of the latent vertex function over the affine family of exact solutions. That latent

function is a scalar perception error rather than a rating, which makes the floor in Proposition I.2 a conditional pointwise-η statement. The general direct theorem uses $\beta \geq \max|e|$ without this projection. We use the graph connection only to motivate a scalar-compatibility diagnostic.

Proposition I.1 (Explainability)**.** *An explaining d exists if and only if every closed walk of H has zero signed residual sum. Consequently, if some closed-walk sum is nonzero, no single scalar perception explains the observed edges* at any η.

Proof. If d explains the observations, the signed contribution of an occurrence traversed from u to v is $d(u) - d(v)$, so the sum along every closed walk telescopes to 0. Conversely, choose a spanning forest using particular edge occurrences, set $\pi = 0$ at one root per component, and propagate π along each chosen occurrence so that its stored equation holds. Consider any unchosen occurrence $q = (a, b, e)$. Traverse q from a to b, then follow the unique tree path from b back to a. The resulting closed walk has signed sum $e + \pi(b) - \pi(a)$, which vanishes by hypothesis. Hence $\pi(a) - \pi(b) = e$. This also covers occurrences parallel to a chosen tree edge; self-loops have value zero by their one-edge closed walks. Thus π explains every occurrence. □

Proposition I.2 (Sharp floor)**.** *Suppose the observations are explainable and let π be a spanning-forest potential. Then*

$$\min\{\, \|d\|_\infty : d \text{ explains } E \,\} \;=\; \max_K \; \frac{\max_{a\in K} \pi(a) - \min_{a\in K} \pi(a)}{2},$$

the maximum running over connected components K of H, with the maximum over an empty family defined as 0*, and the minimum is attained. Thus both sides are* 0 *when there are no observations. In particular η is at least this value, conditional on the shared scalar hypothesis.*

Proof. If d and π both explain the observations then $d - \pi$ has zero difference across every edge, so it is constant on each component: $d = \pi + \theta_K$ on K. Writing $M_K = \max_{a\in K} \pi(a)$ and $\mu_K = \min_{a\in K} \pi(a)$, the quantity $\max_{a\in K} |\pi(a)+\theta_K|$ equals $\max\{M_K+\theta_K,\, -(\mu_K+\theta_K)\}$, which is minimized at $\theta_K = -(M_K + \mu_K)/2$ with value $(M_K - \mu_K)/2$. Components are independent, so taking the maximum over K gives the stated value, realized by that choice of offsets. □

Proposition I.3 (Chaining can be strictly stronger)**.** *If E is nonempty, the floor of Proposition I.2 is at least* $\max_{(a,b,e)\in E} |e|/2$*. The inequality is strict for some observations and tight for others. If the observations in a component consist of a path traversed in one direction whose k occurrences all carry the same residual r, that component's floor is $k|r|/2$, against $|r|/2$ from any single comparison.*

Proof. For an element (a, b, e) inside component K we have $|e| = |\pi(a) - \pi(b)| \leq M_K - \mu_K$, which gives the inequality. On the stated path π changes by $-r$ at each step and is therefore monotone, so $M_K - \mu_K = k|r|$. □

Remark I.4 (The gain is not automatic). Chaining buys nothing when the residuals cancel. A two-element path carrying $e = 1$ then $e = -1$ has $\pi = (0, -1, 0)$, spread 1, and floor $1/2$: exactly the single-comparison bound. What the floor tracks is the *range of the partial sums*, so length helps only to the extent that residuals accumulate in one direction. Any statement that k comparisons give k times the single-comparison bound without a co-orientation hypothesis is false.

The three statements above combine into the following certificate.

Proposition I.5 (Residual-graph certificate)**.** *An explaining d exists if and only if every closed walk has zero signed residual sum. If the observations are explainable and π is a spanning-forest potential, then*

$$\min_{d \text{ explains } E} \|d\|_\infty = \max_K \frac{\max_{a\in K} \pi(a) - \min_{a\in K} \pi(a)}{2}.$$

The maximum is over connected components and is zero for an empty observation set. If E is nonempty, it dominates the single-comparison floor $\max_{(a,b,e)\in E} |e|/2$*. If one connected component consists of a consistently oriented k-edge path with residual r on every edge, that component contributes $k|r|/2$ to the floor.*

Proof. The first sentence is Proposition I.1; the displayed identity and the sentence following it, including the empty-observation convention, are Proposition I.2; and the last two conditional sentences are Proposition I.3, whose proofs are given above with the multigraph edge cases. Remark I.4 is the reason the last sentence carries a co-orientation hypothesis rather than a bare length. □

One shared scalar perception over the comparisons is the added hypothesis. Definition B.1 admits nonintegrable direct reports and time variation. If each comparison carried its own report, the residuals would impose no joint constraint and neither the graph refutation nor its conditional scalar floor would apply. The conclusion therefore splits three ways. The general per-comparison floor $\beta \geq |e|$ always holds. The pointwise floor $\eta \geq |e|/2$ and Proposition I.2 require a scalar explanation, with the graph sharpening additionally requiring one explanation shared across the comparison set. A failure of Proposition I.1 refutes that shared scalar explanation on the edges of E; it does not contradict Theorem C.2. Zero closed-walk sums prove only that the observed multigraph is explainable: they do not certify unobserved pairs, future reports, or stationarity.

Consequences for finite observation families. A family containing one occurrence from a fixed incumbent to each of several distinct nonincumbent candidates, and no other edges, is a star. Its underlying graph is acyclic and, by Proposition I.1, cannot refute a shared scalar explanation. Reversed or repeated occurrences test two-edge consistency; comparisons completing a triangle test additive consistency across three distinct action pairs. A family intended to inspect both kinds of constraint should include both kinds of closed walk. Occurrences can be combined into one graph only when their player, game, other-player state, and reporting context agree. For a forward/reverse pair, a zero sum is antisymmetry; for two occurrences of the same ordered pair, a zero signed sum requires agreement. Neither a triangle check nor a two-edge check alone establishes explainability of an arbitrary observed graph: Proposition I.1 requires zero signed sum on every closed walk.

A minimal diagnostic repeats one ordered pair twice. Doing so already creates a two-edge closed walk, so two residuals that disagree after orientation has been accounted for refute a shared scalar explanation at that pair alone, with no additional states. A repeated-observation discrepancy is often described as noise; here it has the precise consequence above and requires only two edge occurrences.

The scope of these propositions is limited to the observed comparisons and to the stated shared-scalar hypothesis.

J Further scope limitations

The stale-read descent results assume fresh own coordinates, currently legal candidates, simultaneous coordinate writes in one fixed game, and the stated influence budgets. Their age-free specialization is an event-time guarantee; a fresh finite-time endpoint certificate additionally needs bounded read age and either bounded-gap or conditional probabilistic full-scan activation, and its radius needs intrinsically accurate scan reports. With unbounded delay, rejection scans retain only their scan-specific stale radii. The results do not cover arbitrary shared-memory transactions, forgetting, or compression. This paper also does not study how an orchestrator should choose the game, the computational complexity of that choice, or orchestrators that change the game during a run; none of these follows from the results proved here.

K Regime and report-class summaries

Table 1: Termination and the guarantee from an exhaustive rejected scan.

Regime	Universal behavior	Rejected-scan guarantee
$\delta > \beta$	finite, with the bound in Theorem 4.2	$(\delta + \beta)$-NE
$\delta = \beta > 0$	finite, at most $\lvert S\rvert - 1$ moves; no spread rate (Proposition C.3)	2β-NE
$0 < \delta < \beta$	a stationary one-player cycle exists (Proposition C.4)	$(\delta + \beta)$-NE

Table 2: General direct reports and the scalar-representable subclass.

Question	General direct field	Scalar-representable subclass
Primitive	ordered improvement reports	differences of one $\hat{c}_i^u$
Recentering	unavailable	pointwise error $\beta/2$ on a complete fiber
Stationary lone mover	can cycle if $\delta < \beta$	terminates for every $\delta > 0$
Stationary normalized rate	leading constant one is optimal but never attained	one player, $\delta > \beta$: exact asymptotic ceiling $(\delta - \beta)/\delta$
Cycle test	optional scalar diagnostic, not required	nonzero sums refute a shared scalar field; zero sums cover observed edges only

L Optimality of the leading constant

The dependence on $\Delta_\Phi/(\delta - \beta)$ is not an artifact of telescoping. Let $\gamma = \delta - \beta > 0$ and, for an integer $K \geq 1$, consider the single-player game with actions $\{0, 1, \ldots, K\}$ and

$$c(j) = \Phi(j) = (K - j)a_K, \qquad a_K = \gamma \frac{K + 1}{K}.$$

Use the stationary direct report $a_K + \beta$ on every unit-forward pair and the exact true improvement on every other ordered pair. A scheduler that selects the unit step $t \to t + 1$ is legal because $a_K + \beta > \delta$.

Proposition L.1 (Asymptotically tight rate, restated from Proposition C.5)**.** *For every $\beta \geq 0$ and $\delta > \beta$, the stationary direct staircase with actions $\{0, \ldots, K\}$ and $a_K = (\delta - \beta)(K + 1)/K$ has a maximal trajectory with exactly $m = K$ moves and $\Delta_\Phi = (K + 1)(\delta - \beta)$. Hence*

$$\frac{m}{\Delta_\Phi/(\delta - \beta)} = \frac{K}{K + 1} \longrightarrow 1.$$

The leading constant one in Theorem 4.2 is optimal and is never attained by a nonempty finite trajectory.

The proof is in Proposition C.5. For $\beta > 0$, the forward and reverse unit reports do not sum to zero, so this stationary field is not scalar-representable. When $(K - 1)a_K > \beta$, a rule that selects the greatest reported improvement jumps from 0 to K, so there the long path is a scheduler phenomenon. When $K(\delta - \beta) \leq \beta$, every report other than a unit-forward one is at most δ, and every maximal trajectory starting from 0 makes K moves. Thus direct stationarity alone does not improve the leading constant.

Proposition L.2 (Stationary single-player ceiling, restated from Proposition C.6)**.** *Consider a one-player finite exact-potential game with scalar-stationary, scalar-representable, β-accurate reports. For every $\delta > 0$, every nonempty trajectory of m accepted moves satisfies $m < (\Delta_\Phi + \beta)/\delta$. In all cases,*

$$m \leq \max\left\{0, \left\lceil \frac{\Delta_\Phi + \beta}{\delta} \right\rceil - 1\right\},$$

and every maximal trajectory terminates. If $\delta > \beta$, there are scalar-stationary staircase instances with $\Delta_\Phi \to \infty$ for which

$$\frac{m}{\Delta_\Phi/(\delta - \beta)} \longrightarrow \frac{\delta - \beta}{\delta}.$$

Thus this ratio is the exact asymptotic supremum in the stationary single-player scalar class. If $\beta > 0$, it is strictly below the unrestricted leading constant one in Proposition L.1.

The proof and lower family are in Proposition C.6.

This ceiling is scalar-only and is not stated for multiple players: reports along a mixed-player path come from different scalar functions and do not telescope into two endpoints. By contrast, Proposition L.1 already shows that unrestricted stationary direct reports approach, without attaining, the general leading constant one.

Proposition L.3 (Robust local-to-global alignment, restated)**.** *Proposition 3.2: for a fixed player i and fixed opponent utilities, every strict unilateral w_i-improvement by player i strictly improves the total utility Ω for every realization of w_i if and only if the aggregate Ω_{-i} does not depend on player i's action.*

Proof. For a unilateral deviation write $\Delta\Omega = \Delta w_i + \Delta\Omega_{-i}$. If the other players' aggregate is independent of s_i, then $\Delta\Omega_{-i} = 0$ and every positive Δw_i is a positive $\Delta\Omega$. Conversely, suppose that independence fails. After swapping the two actions if necessary, there are a, b, s_{-i} with $D = \Omega_{-i}(a, s_{-i}) - \Omega_{-i}(b, s_{-i}) > 0$. Choose $0 < \nu < D$ and define $w_i(b, s_{-i}) = \nu$ and $w_i = 0$ elsewhere. The deviation $a \to b$ improves w_i by ν but changes the global objective by $\nu - D < 0$. □

M AN ILLUSTRATION WITH RECORDED LANGUAGE-MODEL REPORTS

This section applies the record-level checks of Section 7 to comparisons reported by one language model in two small workboard games whose true gains are exactly computable. It shows how those checks read on recorded reports and which of their premises fail. No move was executed, so nothing here concerns trajectories or convergence, and the records are too few and too specific to estimate the model's error.

Setup. A *game* is a workboard game of Proposition 3.1 with workers 0 and 1 (indexed from 0, as in the prompts) and tags 0, 1 and 2, in which each worker holds exactly one tag from this common set. All payout schedules are nonincreasing, so each coverage function is concave. A *profile* lists both workers' tags, so $(2, 1)$ means that worker 0 holds tag 2 and worker 1 holds tag 1. A *directed comparison* fixes a game, a profile, a worker and one alternative tag for that worker. It is *beneficial* when its true gain is positive and *harmful* when it is negative; every queried true gain is a nonzero integer. A *query* presents one directed comparison to the model, and a *response* is one reply to one query. The *complete scan* of a profile consists of its four directed comparisons, two for each worker. Table 3 lists the two games.

Table 3: The two games. Each entry gives the amount paid to every holder of the tag when one or two workers hold it, that is, the marginal credits $r_\tau(1)$ and $r_\tau(2)$ of Proposition 3.1.

Game	Tag 0	Tag 1	Tag 2	Best coverage	Pure equilibria
G1	$7, 4$	$4, 1$	$8, 7$	15	$(0, 2)$, $(2, 0)$, $(2, 2)$
G2	$3, 2$	$7, 4$	$6, 4$	13	$(1, 2)$, $(2, 1)$

Queries and model. Each query is a single request, one system and one user message, with no conversation history. It gives the payout rule, the game's schedules, both workers' current tags, and the queried worker's current tag and alternative, each explicitly labeled. It asks for a JSON object with three fields: the worker's estimated payout under the alternative minus its current payout; an action, `move` exactly when that estimate is strictly greater than 0.5 and `stay` otherwise; and a reason of under twenty words. The estimate is the report $\widehat{\Delta}$ of Definition 2.1 with margin $\delta = 1/2$. The true gain Δ is the utility difference, which equals the change in coverage by Proposition 3.1. All prompts are synthetic. We requested `openai/gpt-5.2` through OpenRouter with routing restricted to Azure, and the retained metadata give the endpoint identifier `openai/gpt-5.2-20251211`. Each request set seed 20260907, reasoning effort `none` and a completion cap of 512 tokens. Temperature and top-p were not sent, so their effective values are unknown, and repeats with identical messages and seed sometimes returned different numbers. Model names are recorded fields; the serving endpoint was not independently authenticated.

Two designs. The two designs were collected separately and are never pooled or compared. Both followed four exploratory attempts, one of which completed a smaller panel. The collection protocol changed during those attempts, and no response from them is reported or pooled here. *Design I* (64 responses) asks eight directed comparisons of worker 0 while worker 1 holds tag 1: in each game, both directions between tags 2 and 0 and between tags 2 and 1. Each is posed with each of two context passages, a neutral note about system logs or an invented workboard log recording an agreement to

keep coverage spread out, and with the two options listed in either order. The resulting 32 queries were each asked twice, the second pass taking them in reverse sequence. Of the 64 responses, 32 concern beneficial and 32 harmful comparisons. *Design II* (16 responses) asks the complete scan of profile $(2, 1)$ in each game, with the neutral passage and one option order, twice; 4 responses concern beneficial and 12 harmful comparisons. Design II was chosen after inspecting the Design I responses. Each design began with two format-check requests, fixed in advance and excluded from every count.

What is checked. For each response the residual is $e = \widehat{\Delta} - \Delta$, as in Section 7, and the report is *exact* when $e = 0$. A response *nominates* its alternative when $\widehat{\Delta} > 1/2$. Its number-implied decision is wrong when it nominates a harmful comparison or fails to nominate a beneficial one, and its stated action is wrong in the same sense. For a nomination, $e < \delta$ is the per-move condition of Proposition 7.1: had the move been made, coverage would have risen. The stronger condition $|e| < \delta$ is the positive-gap condition of Proposition D.1 with the smallest radius the record allows, $\beta_i^u(s, a) = |e|$. A complete scan certifies its profile through Proposition 7.1 only if it nominates nothing. For two reversed comparisons, from a to b and from b to a against the same tag of the other worker, the true gains cancel. A scalar explanation shared by both reports therefore requires the reports to sum to zero, the two-edge case of Proposition I.5.

Design I. Table 4 gives the counts.

Table 4: Design I: 64 responses to eight directed comparisons.

Exact numerical reports	44
Nominations: beneficial / harmful	32 / 9
Wrong number-implied decisions	9
Wrong stated actions	12
Stated action contradicting the response's own number	3
Nominations with $e < \delta$	32
Nominations with $\lvert e\rvert < \delta$	24
Largest and mean $\lvert e\rvert$	4 and $21/32$
Repeated queries whose number changed	7 of 32
Reversed pairs with nonzero report sum	16 of 32

Design II. Table 5 lists the four complete scans. Across its 16 responses there are 3 exact numbers, 5 wrong number-implied decisions, 6 wrong stated actions, 1 stated action contradicting its own number, and 2 nominations with $e < \delta$. The largest $|e|$ is 5 and the mean is $43/16$, and 4 of the 8 repeated queries changed their number.

Table 5: Design II: complete scans of profile $(2, 1)$. Regret is the smallest $\varepsilon \geq 0$ for which the profile is an ε-Nash equilibrium, that is, the larger of 0 and the largest true gain available to either worker.

Scan	Regret	Nominations: beneficial / harmful	Every candidate rejected
G1, first	3	2 / 0	no
G1, second	3	2 / 0	no
G2, first	0	0 / 3	no
G2, second	0	0 / 2	no

What the records show. Figure 1 places every response of each design at its true gain and report. Each observed residual gives only a lower bound, $\beta \geq |e|$. The largest residuals, 4 in Design I and 5 in Design II, all come from over-reporting a harmful move at profile $(2, 1)$ of game G2. Any error bound that covers these records is at least 5, ten times the margin, so the termination guarantee of Theorem 4.2, which needs $\delta > \beta$, cannot be invoked at this margin. Both per-move conditions fail on every harmful nomination by construction, since a nomination with $e < \delta$ has positive true gain. They differ on beneficial nominations: in Design I all 32 satisfy $e < \delta$, while only 24 satisfy $|e| < \delta$, because eight responses report 6 for a true gain of 7. The signed condition is sufficient but not necessary for a beneficial move: in Design II two beneficial nominations report 6 for a true gain

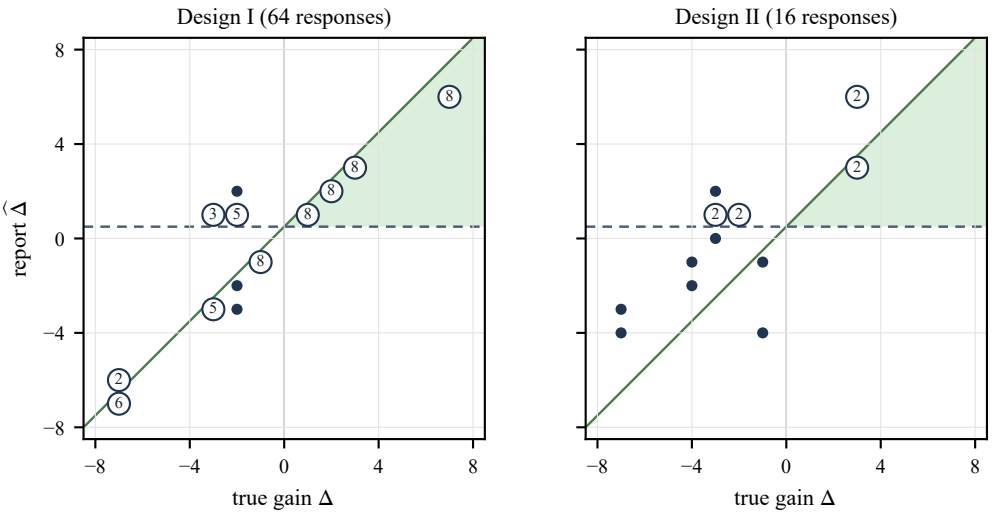

Figure 1: The 64 responses of Design I (left) and the 16 of Design II (right), each at its true gain Δ and report $\widehat{\Delta}$; a numeral gives the number of responses at a point, and a dot marks a single response. The dashed line is $\widehat{\Delta} = \delta$ and the solid line is $\widehat{\Delta} = \Delta + \delta$, with $\delta = 1/2$. The shaded region between them, boundaries excluded, is $\widehat{\Delta} > \delta$ with $e < \delta$, the per-move condition of Proposition 7.1, so every point in it has $\Delta > 0$; it holds 32 of the 41 Design I nominations and 2 of the 9 Design II nominations. Stated actions are not shown, and the two designs are not compared.

of 3 and fail it. Applying higher margins to the same recorded Design I numbers accepts 24 beneficial and 1 harmful response at margin 1, and 16 and 0 at margin 2, against 32 and 9 at $1/2$. The prompts stated the margin 0.5, so this says nothing about reports elicited under another margin or about other records.

Reports that no single scalar explains. Pair each Design I response with the response to the reversed comparison under the same context passage, option order and pass. Of these 32 pairs, 16 have reports with a nonzero sum, and any other pairing gives the same count, because within each comparison the reports in the beneficial direction are all equal. So no single scalar explains both reports of such a pair. This rules out one scalar shared by the two occurrences of a pair, not different scalars at different occurrences, which Definition 2.2 permits. Likewise, each of the 7 repeated Design I queries whose number changed admits no scalar shared by its two occurrences.

Actions and numbers. The stated action contradicts the response's own number in 3 of 64 and 1 of 16 responses, although the query defines the action from the number; in every reply the action field precedes the number. The checks above therefore apply the acceptance rule of Definition 2.1 to the number and treat the stated action as a separate channel.

Complete scans. None of the four complete scans in Design II rejects every candidate, so the premise of the report-based terminal certificate in Proposition 7.1 is never met. An unmet premise does not mean that the profile is not an equilibrium. At $(2, 1)$ in G1, which is not one (regret 3), each scan nominated both improving moves. At $(2, 1)$ in G2, a pure equilibrium with maximal coverage 13, the game satisfies the hypotheses of Proposition 5.2, under which every pure equilibrium has maximal coverage. Yet the two scans there nominated three and two harmful moves, so a system acting on these reports at $\delta = 1/2$ would leave an optimal equilibrium.

One comparison. This Design I comparison was chosen after inspecting the responses. In game G2 at profile $(2, 1)$, worker 0 considers tag 1. Its payout would fall from 6 to 4, a true gain of -2, and coverage would fall from 13 to 11. One response reported $+2$ with action `move`; the same query, asked again in the second pass, reported -2 with action `stay`. The records do not show why. Reasons were recorded but not scored, and the first response's reason is itself wrong.

Limits. The records support statements about recorded comparisons at fixed profiles and nothing more. No move was executed, so they say nothing about trajectories, convergence or absorption. They give lower bounds on error, not a uniform or future bound, and estimate neither β nor η. The alternatives were enumerated by the queries, not proposed by the model. Repeated queries are not independent samples, and the two designs differ in construction and are not compared. In Design I every move away from tag 2 is harmful and every move to it beneficial, so gain sign and direction are confounded; in Design II the sign is confounded with the worker and the game. Replies were schema-constrained with reasoning effort `none`, and no free-form or reasoning variant was run, so format effects and arithmetic difficulty are not separated. Every prompt labels the current and alternative options explicitly, so option order is a weak manipulation. No queried true gain equals 0 or $1/2$ and no report equals $1/2$, so the margin boundary itself is untested. No effect of the context passages is claimed, nothing here concerns other models, and nothing describes the deployed workboard of Huang et al. (2026).

Records. All counts are exact and were computed from the recorded responses. The supplementary archive provides prompts for the primary comparisons and anonymized response records. Its offline replay code reproduces the numerical summaries and figure from these records.