A RATE SPECTRUM FOR MOVING PROXIMAL CENTERS IN SEQUENCE-FORM GAMES

Anonymous authors
Paper under double-blind review

ABSTRACT

Modern solvers for two-player zero-sum extensive-form games deploy the current iterate as the policy, so the last-iterate question is when it converges. In the regularized counterfactual regret family, convergence comes from a reference strategy. The iterate is regularized toward a reference, which removes the cycling of unregularized dynamics at the price of pulling the solution away from the Nash set. On the Euclidean carrier a fixed magnet converges linearly to a regularized equilibrium at distance $O(\tau)$ from the Nash set, for proximal weight τ. Periodic re-centering and hard resets reach exact equilibria at asymptotic or sublinear rates. Between these ends lies the interior speed $\beta_t \in (0, 1)$, for which no finite-time rate is known for this update class. All three move one reference at different speeds, so their common-carrier proxies are instances of one schedule coordinate, the prox-center update spectrum, and one analysis covers the carrier-matched family. A constant speed inside an explicit window gives linear last-iterate convergence to an exact Nash equilibrium of the original game; the bound is sharp at the endpoint. Summable schedules freeze at the computable resolvent of the center's limit, with a closed-form positive floor at interior equilibria. At interior equilibria the asymptotic slope is a closed form in the speed, the proximal weight, and the game's skew spectrum, and its mode eigenvalues match measured slopes to three significant figures. On Leduc, whose equilibrium set is a manifold, every practical speed lies outside the estimated window and the finite-horizon decay at the faster speeds is fitted by a power law. Linear decay reaches a saddle gap of 10^{-10} on Kuhn, and the Leduc power-law exponent is 1.8–2.2. Across twenty-eight carrier-matched entries the moving center reaches a saddle gap below 10^{-6} wherever any other method does. On Leduc it leads periodic re-centering by 2–3$\times$ at 500k iterations.

1 INTRODUCTION

Counterfactual regret minimization (CFR) (Zinkevich et al., 2007) and its proximal descendants are the workhorse solvers for two-player zero-sum extensive-form games. Their classical guarantee covers the *average* strategy, the policy that the poker milestones deployed (Bowling et al., 2015). A recent line asks the *last* iterate itself to converge (Lee et al., 2021).

In the regularized family the mechanism that delivers last-iterate convergence is a reference strategy. The iterate is regularized toward a reference, which removes the cycling of the unregularized dynamics at the price of moving the solution away from the Nash set (Sokota et al., 2023). Magnetic mirror descent (MMD) keeps a fixed magnet (Sokota et al., 2023) and converges linearly to a *regularized* equilibrium. On the Euclidean carrier that point sits at distance $O(\tau)$ from the Nash set, τ the proximal weight. Adaptively perturbed mirror descent (APMD) re-centers periodically (Abe et al., 2024), and Reward Transformation CFR$^+$ (RTCFR$^+$) and its adaptive variant reset an outer reference on a schedule or on an exploitability trigger (Meng et al., 2025; Ren et al., 2025). Both reach exact equilibria at asymptotic or sublinear rates. A separate line uses optimism with no reference strategy at all (Daskalakis & Panageas, 2019; Wei et al., 2021; Cai et al., 2024).

Each method tunes the same quantity: how fast the reference follows the iterate. A re-centering weight moves the center at an interior speed $\beta_t \in (0, 1)$; a hard reset is the endpoint $\beta_t = 1$. The interior case, one projected step per center update on the treeplex, has, for this update class, no known finite-time rate to an exact equilibrium of the original game. No criterion locates the

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.

boundary between a center too slow to remove the bias and one too fast for the iterate to track. We ask when a moving reference converges, how fast, and to what.

We answer these questions on one coordinate. Each mechanism above has a common-carrier proxy that updates its reference by $\mu_{t+1} = (1 - \beta_t)\mu_t + \beta_t z_{t+1}$, where z_{t+1} is the iterate and μ_t the reference (Section 3). On that carrier the update is a β-relaxed *inexact* proximal-point iteration, bounded by an explicit 2×2 matrix (Section 4). We call the family the prox-center update spectrum, an analysis coordinate on which the theorems below hold for every carrier-matched member at once. The native algorithms enter only through these proxies.

Experiments test each claim where it is sharpest. On Kuhn, whose equilibrium set is non-singleton, every constant schedule inside the estimated window decays linearly with no floor, reaching a saddle gap of 10^{-10}. Across twenty-eight carrier-matched entries the moving center reaches a saddle gap below 10^{-6} wherever any other common-carrier method does. On Leduc it trails periodic re-centering by 9% at 200k but leads it by 2–3$\times$ at 500k (Section 8). Leduc also marks the limit of the guarantee. Its equilibrium set is a manifold and its inner contraction factor is $\rho_{\text{in}} \approx 0.95$. Every practical center speed therefore sits outside the estimated window, and the finite-horizon behavior is power-law-like with exponent $\alpha \approx 1.8$–2.2. Three structural interventions probe that breakdown (Section 7).

Our contributions:

- We prove that constant speeds inside the explicit window Equation (9) give last-iterate linear convergence to an *exact* Nash equilibrium of the original game. The window endpoint is sharp for the two-dimensional bound (Theorem 4.5).
- We characterize three further schedule classes. Summable schedules freeze at the computable resolvent of the center's limit, in general a non-equilibrium, with a closed-form positive floor at interior equilibria the center cannot reach, where the asymptotic slope is a closed form in the speed, the proximal weight, and the game's skew spectrum (Theorem 5.1 and Proposition B.1). Nonincreasing in-window schedules inherit the exponent, and vanishing speeds with divergent mass converge at a power-law rate for harmonic decay (Theorems 6.1 to 6.3).
- We delimit the domain of validity on four games and ten methods, twenty-eight entries on the common carrier. The behavior is linear on Kuhn, power-law-like on Leduc over the horizon, and unchanged under three structural interventions (Sections 7 and 8).

2 Related Work

Last-iterate convergence in zero-sum and extensive-form games rests on forgetful dynamics (Daskalakis & Panageas, 2019; Wei et al., 2021; Lee et al., 2021; Cai et al., 2024); on the optimization side the same re-centering appears as a moving anchor (Alcala et al., 2025b; Chen, 2026), analyzed under monotone or contractive structure. The two ends of Section 1 are the fixed magnet of MMD (Sokota et al., 2023) and the resets of APMD and RTCFR$^+$ (Abe et al., 2024; Meng et al., 2025); perturbing one player's payoff alone reaches an equilibrium of the unperturbed game at a linear rate on bilinear problems (Abe et al., 2026). Closest in mechanism are the drifting proximal center of Guo & Han (2020) and the interpolation between extragradient and proximal steps of Pethick et al. (2025) for weak monotone variational inequalities: neither yields a finite-time, rate-explicit statement for $\beta_t \in (0, 1)$, which Theorem 4.5 does by absorbing the tracking error into a coupled two-state system. Appendix H surveys the CFR line and the remaining anchor and reference-update literature.

3 Preliminaries

Zero-sum extensive-form games in sequence form We study two-player zero-sum extensive-form games with perfect recall in their sequence-form representation, the setting of CFR and its descendants (Zinkevich et al., 2007). Each strategy space is a treeplex $X = \{x \geq 0 : Ex = e\}$ cut out by the sequence-form flow constraints (E, e), the polytope on which regret matching is an instance of mirror descent (Liu et al., 2022; Farina et al., 2019). The joint space is $Z = X \times Y$ with

$z = (x, y)$. The payoff matrix A defines the bilinear saddle problem and the operator

$$F(z) \;=\; (-Ay,\; A^\top x), \tag{1}$$

which is linear, skew, monotone, and L-Lipschitz with $L = \|A\|_2$, the constrained monotone setting of the last-iterate literature (Daskalakis & Panageas, 2019; Wei et al., 2021). The zero set of $F + N_Z$ is the equilibrium set $Z^\star = \{z \in Z : 0 \in F(z) + N_Z(z)\}$, with N_Z the normal cone of Z. Solution quality is measured by the sequence-form saddle gap, twice the exploitability standard that evaluators such as OpenSpiel report (Lanctot et al., 2019; Bowling et al., 2015; Lee et al., 2021),

$$\mathrm{NashConv}(z) \;=\; \max_{x' \in X} x'^\top A y \;-\; \min_{y' \in Y} x^\top A y', \tag{2}$$

which vanishes exactly on $Z^\star$ and is Lipschitz with constant L_{gap} (Proposition 4.4). All statements and the experiments of Sections 7 and 8 evaluate the *last* iterate under $\mathrm{NashConv}$; the rate checks of Section 6 also report $\mathrm{dist}(z_t, Z^\star)$.

Throughout, the *carrier* of a method is the geometry its update runs in; ours is the Euclidean Z of Equation (1). A method is *carrier-matched* when its update takes the form Equation (4).

The proximal step and the prox-center update spectrum The inner update is a single projected proximal step regularized toward a point $\mu \in Z$, the *proximal center*, the magnet of MMD (Sokota et al., 2023) in Euclidean geometry:

$$G_\mu(z) \;\coloneqq\; \Pi_Z\left[\frac{z + \eta\tau\mu - \eta F(z)}{1 + \eta\tau}\right], \tag{3}$$

with step size $\eta > 0$ and proximal weight $\tau > 0$. The full dynamics couple the iterate with a moving center:

$$z_{t+1} \;=\; G_{\mu_t}(z_t), \qquad \mu_{t+1} = (1 - \beta_t)\,\mu_t + \beta_t\,\sigma_t, \qquad \sigma_t = z_{t+1}, \tag{4}$$

indexed by the schedule $(\beta_t)_t$ with $\beta_t \in [0, 1]$ throughout; we call this family the *prox-center update spectrum*. Its endpoints and interior recover the production reference-strategy mechanisms (Section 2): $\beta_t \equiv 0$ fixes the center as in MMD (Sokota et al., 2023), $\beta_t \in \{0, 1\}$ resets it by hard jumps as in APMD and RTCFR$^+$ (Abe et al., 2024; Meng et al., 2025; Ren et al., 2025), and $\beta_t \in (0, 1)$ moves it continuously, the interior studied here. Viewed on the center, Equation (4) is a relaxed inexact proximal-point iteration (Pethick et al., 2025).

Two scalars track the coupled system Equation (4). Under Lemma 4.1, G_μ is a contraction with unique fixed point $z^\star_\tau(\mu)$, and we write

$$e_t \;\coloneqq\; \|z_t - z^\star_\tau(\mu_t)\| \qquad \text{(tracking error)}, \qquad a_t \;\coloneqq\; \mathrm{dist}(\mu_t, Z^\star) \qquad \text{(center error)}. \tag{5}$$

The rates of the paper are expressed through two spectral quantities: the inner contraction factor ρ_{in} of G_μ, and the contraction factor ρ_J of the fixed point map toward $Z^\star$, which comes from an error bound of metric subregularity type (Robinson, 1981). Both are given in closed form in Section 4, where the coupled evolution of (e_t, a_t) is bounded by the 2×2 tracking matrix $M(\beta)$. The dynamics take a single inner step per outer update and carry the solver state across steps, a *warm start* (Section 4).

4 The Window

This section proves the central positive result. When the center moves at constant speed β inside an explicit window, the last iterate of Equation (4) converges linearly to an exact Nash equilibrium of the *original* game, with no bias floor. A fixed center, by contrast, carries an $O(\tau)$ offset (game and center fixed, $\tau \to 0$). Three lemmas supply the spectral data $(\rho_{\mathrm{in}}, \rho_J)$, a two-dimensional recursion, and an endpoint-sharp spectral condition (proofs in Appendix A).

Lemma 4.1 (Inner contraction)**.** *For every $\mu \in Z$, the map G_μ of Equation* (3) *is ρ_{in}-Lipschitz with*

$$\rho_{\mathrm{in}} \;=\; \frac{\sqrt{1 + \eta^2 L^2}}{1 + \eta\tau}, \qquad \textit{and} \qquad \rho_{\mathrm{in}} < 1 \iff \eta\,(L^2 - \tau^2) < 2\tau. \tag{6}$$

Under this step-size condition G_μ has a unique fixed point $z^\star_\tau(\mu)$, with contraction factor ρ^t_{in}.

The inequality follows from skewness and nonexpansiveness; the step-size condition is sufficient.

Lemma 4.2 (The fixed point is a resolvent)**.** $z_\tau^\star(\mu) = J(\mu) \coloneqq (I + \tau^{-1}(F + N_Z))^{-1}(\mu)$*, the resolvent of the maximally monotone operator* $F + N_Z$ *with step* $1/\tau$*. Consequently* J *is firmly nonexpansive,* $\mathrm{Fix}(J) = Z^\star$*, and* $\mu \mapsto z_\tau^\star(\mu)$ *is nonexpansive (Bauschke & Combettes, 2020).*

Writing $\mu_{t+1} = (1-\beta)\mu_t + \beta J(\mu_t) + \beta(z_{t+1} - J(\mu_t))$ exhibits the center dynamics as a β-relaxed *inexact* proximal-point iteration. Its inexactness is the tracking error of a single warm-started inner step. The certificate-based analyses of Pethick et al. (2025) treat that error separately; here it is absorbed into the coupled system.

Lemma 4.3 (Resolvent contraction toward $Z^\star$)**.** (a) *Suppose* $T = F + N_Z$ *is metrically subregular on* $Z^\star$ *with modulus* γ*, i.e.* $\mathrm{dist}(z, Z^\star) \le \gamma \cdot \mathrm{dist}(0, T(z))$ *on* Z*. Then*

$$\mathrm{dist}(J(\mu), Z^\star) \;\le\; \rho_J \cdot \mathrm{dist}(\mu, Z^\star), \qquad \rho_J \;=\; \frac{\gamma\tau}{\sqrt{1+\gamma^2\tau^2}} \;<\; 1. \tag{7}$$

Polyhedrality gives subregularity (Robinson, 1981), with uniform modulus along the trajectory. (b) *If the equilibrium is unique and relatively interior with tangent space* $V \neq \{0\}$*, then locally* $\rho_J = \tau/\sqrt{\tau^2 + \lambda^2}$*, where* λ *is the smallest nonzero eigenvalue modulus of the compression* $P_V F P_V$ *of the skew map to* V *(equivalently* $\gamma = 1/\lambda$*);* $V = \{0\}$ *gives* $\rho_J = 0$*.*

Part (b) is exact where it applies; on boundary or non-unique equilibria only part (a) applies, and the modulus γ becomes a measurable quantity exploited in Section 7 (numeric checks in Appendix A).

Proposition 4.4 (Two-dimensional tracking recursion)**.** *Under the conditions of Lemma 4.1 and Lemma 4.3(a), for every* $\beta_t \in [0, 1]$*,*

$$\begin{aligned} e_{t+1} &\le \rho_{\mathrm{in}}(1+\beta_t)\, e_t \;+\; \beta_t\, a_t, \\ a_{t+1} &\le \big(1 - \beta_t(1-\rho_J)\big)\, a_t \;+\; \beta_t \rho_{\mathrm{in}}\, e_t, \end{aligned} \qquad \textit{i.e.} \qquad [e;a]_{t+1} \;\le\; M(\beta_t)[e;a]_t, \tag{8}$$

elementwise, with $M(\beta) = \begin{bmatrix} \rho_{\mathrm{in}}(1+\beta) & \beta \\ \beta\rho_{\mathrm{in}} & 1 - \beta(1-\rho_J) \end{bmatrix}$*. Moreover* $\mathrm{dist}(z_t, Z^\star) \le e_t + \rho_J\, a_t$*, hence* $\mathrm{NashConv}(z_t) \le L_{\mathrm{gap}}(e_t + \rho_J a_t)$*, with* $L_{\mathrm{gap}} \le \sqrt{2}\, L\, R$ *(*R *the strategy-set radius;* $R = 1$ *for simplex games).*

The two rows formalize the tension. Larger β pulls a_t down at speed $\beta(1-\rho_J)$, but it injects $\beta(\rho_{\mathrm{in}} e_t + a_t)$ into the tracking error, whose contraction budget is only $1 - \rho_{\mathrm{in}}$.

Theorem 4.5 (Windowed linear convergence to exact Nash equilibria)**.** *Assume* $\eta(L^2 - \tau^2) < 2\tau$ *and let* $\beta \in (0, 1)$ *be constant with*

$$\beta \;<\; \bar\beta \;\coloneqq\; \frac{(1-\rho_{\mathrm{in}})(1-\rho_J)}{\rho_{\mathrm{in}}\,(2-\rho_J)}. \tag{9}$$

Then the spectral radius satisfies $\rho(M(\beta)) < 1$ *and, for every* t*,*

$$\mathrm{NashConv}(z_t) \;\le\; L_{\mathrm{gap}} \sqrt{1+\rho_J^2}\; \rho_{\mathrm{in}}^{-1/2}\, \rho(M(\beta))^{\,t}\, \big\|[e_0;\, a_0]\big\|, \tag{10}$$

so the last iterate converges linearly to an exact Nash equilibrium of the original game. The endpoint is sharp: at $\beta = \bar\beta$ *(when* $\bar\beta \le 1$*, as in every game here) the boundary condition* $1 - \mathrm{tr}\, M(\beta) + \det M(\beta) = 0$ *for the* 2×2 *nonnegative matrix* $M(\beta)$ *is Equation* (9) *with equality, so* $\rho(M(\beta)) = 1$*.*

Sharpness is a property of the comparison matrix $M(\beta)$. The true iterates can converge linearly beyond $\bar\beta$, as Kuhn does (Section 8), and can also lose local stability (Appendix C). The inner factor ρ_{in} is attained on the two matrix games, where the measured worst-case contraction ratios 0.979121 and 0.962091 equal ρ_{in} to six digits; it is conservative on Kuhn, 0.909640 against $\rho_{\mathrm{in}} = 0.917$. The slack has an explicit source: $\|F\delta\| = L\|\delta\|$ holds only along the top singular directions of A, which span the simplex tangent space of the matrix games and need not be feasible directions of a treeplex.

Perturbing the slow eigenvalue gives the small-β expansion

$$\rho(M(\beta)) \;=\; 1 - \beta(1-\rho_J) + O\Big(\tfrac{\beta^2}{1-\rho_{\mathrm{in}}}\Big) \qquad \Longrightarrow \qquad \text{decay of the bound} \approx\; \beta\,(1-\rho_J), \tag{11}$$

for a fixed game and (η, τ) as $\beta/(1-\rho_{\mathrm{in}}) \to 0$. The measured slope of $\log \mathrm{NashConv}$ is compared with this coefficient in Section 6. The recursion itself is verified step by step with zero violations over three matrix games at $\beta \in \{0.01, 0.05, 0.2\}$ and on the Kuhn treeplex at $\beta \in \{0.002, 0.01, 0.05\}$. The certificate is worst-case (Section 8); sharp local rates are Section 6.

Time-varying schedules enter the guarantee through a persistent excitation condition.

Definition 4.6 (Windowed cumulative excitation)**.** A schedule $(\beta_t)_t$ has windowed cumulative excitation if there exist $K \in \mathbb{N}$ and $\varepsilon > 0$ such that $\sum_{s=t}^{t+K-1} \beta_s \geq \varepsilon$ for each t.

A pointwise floor satisfies Definition 4.6 with $K = 1$; capped below $\bar{\beta}$, the property gives linear convergence and divergent total mass alone gives convergence (Corollary A.1). Speeds with $\beta_t \to 0$ violate the definition; under the hypotheses of Theorem 6.3(ii), divergent cumulative mass reaches the equilibrium without a geometric rate, given local entry with a bounded tracking-to-center error ratio on every skew mode and the $\lambda_{\min}$ mode excited.

5 Summable Schedules

Theorem 4.5 says how fast a well-excited center converges; this section characterizes the opposite end, schedules with finite total movement. The center stops at μ_∞; the iterate converges to the resolvent of that stopping point and $\mathrm{NashConv}$ settles at a computable value. That limit is a non-equilibrium whenever the resolvent leaves the Nash set, and it is lower-bounded in the interior case. Proofs are in Appendix B.

Theorem 5.1 (Freezing under summable schedules)**.** *Under the step-size condition $\eta(L^2-\tau^2) < 2\tau$ (Lemma 4.1), let $(\beta_t)_t$ be any schedule in $[0,1]$ with $S \coloneqq \sum_t \beta_t < \infty$, and let $D = \mathrm{diam}(Z)$. Then (i) $\mu_t \to \mu_\infty$ with $\|\mu_\infty - \mu_0\| \leq SD$; and (ii) $z_t \to J(\mu_\infty)$, in general without a linear rate, and* $\lim_t \mathrm{NashConv}(z_t) = \mathrm{NashConv}(J(\mu_\infty))$.

Part (i) is a Cauchy argument; part (ii) runs the tracking error through a stable convolution that vanishes because $\sum_k \beta_k < \infty$. Theorem 5.1 assumes no error bound.

Freezing produces a positive floor *if and only if* $J(\mu_\infty) \notin Z^\star$. A non-equilibrium stopping point can still resolve into $Z^\star$: at a boundary equilibrium the resolvent maps $\mu_\infty \notin Z^\star$ directly onto $Z^\star$. Explicit floors therefore live in the interior case, where they are computable in closed form (Proposition B.1 and Example 5.2).

Example 5.2 (Matching pennies, fully explicit)**.** For $A = \left[\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right]$ in centered coordinates the resolvent is linear near the equilibrium (wherever the unconstrained solution stays in the box), giving $\mathrm{NashConv} \geq \tau/\sqrt{\tau^2+4}\,\|\mu_\infty\|$ there (with $\mathrm{NashConv} = |u^\star| + |v^\star|$ in closed form, Appendix B); frozen runs land on $\mathrm{NashConv}(J(\mu_\infty))$ digit for digit (0.43584, 0.29803, 0.16077). On Kuhn, $\beta_t = 0.04/(1+t)^2$ freezes at $\mathrm{NashConv} = 3.4485 \times 10^{-1} = \mathrm{NashConv}(J(\mu_\infty))$, with center displacement 0.0149.

An excitation dichotomy Theorem 5.1 covers $\sum_t \beta_t < \infty$. The complementary condition $\sum_t \beta_t = \infty$ might be expected to restore geometric convergence; under the hypotheses of Theorem 6.3(ii) a vanishing speed with divergent mass instead reaches the exact equilibrium with logarithmic exponent $(1-\rho_J^2)\sum_{s<t}\beta_s$, the power law $t^{-c(1-\rho_J^2)}$ for $\beta_t = c/(1+t)$. The experiment shows the consequence (Figure 2 and Appendix G). The summable arms freeze at $\mathrm{NashConv}(J(\mu_\infty))$ near the game scale, the windowed-excitation arms converge geometrically, and the harmonic arm, which keeps divergent total mass, ends nine to ten orders above them and falls on the frozen side of both protocol gates, minimum window-500 sum and tail slope. Over this horizon windowed excitation separates the seven arms of Figure 2, while total mass alone does not. These gates test the finite horizon only: a constant speed above the window satisfies Definition 4.6 with $K = 1$ and can diverge.

6 Local Rate Spectrum

At an interior equilibrium the worst-case constants of Theorem 4.5 sharpen to a closed form. The first-order coefficient of the asymptotic slope depends on (β, τ) and the game's skew spectrum, with no constant to fit. Proofs are in Appendix C.

Theorem 6.1 (Interior local rate)**.** *Suppose the equilibrium is unique and relatively interior, the step-size condition $\eta(L^2 - \tau^2) < 2\tau$ holds (Lemma 4.1), and let $\lambda_{\min} > 0$ be the smallest nonzero skew eigenvalue modulus of the tangential compression of F. Then in a neighborhood of the equilibrium the iteration Equation* (4) *is linear and its asymptotic rate decomposes across the skew modes λ_k of F. The slow mode satisfies*

$$\mathrm{rate}(\beta) \;=\; 1 \;-\; \beta\,(1-\rho_J^2) \;+\; O(\beta^2), \qquad \rho_J^2 = \frac{\tau^2}{\tau^2 + \lambda_{\min}^2}, \tag{12}$$

as $\beta \to 0$ with the game and (η, τ) fixed, so the slope of $\log \mathrm{NashConv}$ *has first-order coefficient $\beta\,\lambda_{\min}^2/(\tau^2 + \lambda_{\min}^2)$. A given trajectory's rate is the dominant modulus among the eigenmodes it excites, equal to the slow-mode modulus when the $\lambda_{\min}$ slow eigenvector is excited; below a threshold $\beta_{\mathrm{s}} > 0$, depending on the game and (η, τ), the $\lambda_{\min}$ slow eigenvalue dominates every eigenvalue of every mode, and the worst-case local rate over initializations is governed by $\lambda_{\min}$.*

The proof diagonalizes the coupled system over the skew eigenpairs $\pm i\lambda_k$ of the tangent space, replacing ρ_J by ρ_J^2 in Equation (11).

Equation Equation (12) reproduces the measured slopes without fitted constants where its hypotheses hold. At $\beta = 0.01$ the first-order coefficient predicts 9.41×10^{-3} on matching pennies and 9.23×10^{-3} on rock–paper–scissors, against measured slopes of 9.51×10^{-3} and 9.34×10^{-3} for $\log\|z_t - z^\star\|$ (fit records in Appendix D). The exact slow eigenvalue of the 2×2 mode system in Appendix C gives 9.513×10^{-3} and 9.341×10^{-3}, matching the measurements to three significant figures. Outside the hypotheses, a random 3×4 game with a unique boundary equilibrium and $\rho_J = 0.665$ predicts 5.58×10^{-3} against a measured 5.65×10^{-3} (same observable and window).

Extrapolating beyond small β gives a heuristic shape prediction. If the center-limited slope $\beta(1-\rho_J^2)$ is capped at the inner budget $1 - \rho_{\mathrm{in}}$, the two lines intersect at

$$\beta^\star \;\approx\; \frac{1-\rho_{\mathrm{in}}}{1-\rho_J^2}. \tag{13}$$

The curve grows linearly below the crossover and degrades *non-monotonically* above it (measured values in Appendix C). Equation Equation (13) extrapolates the small-β expansion. On matching pennies the fast mode of the 2×2 mode system overtakes the falling slow mode and rises above one past the crossover (moduli in Appendix C); at weak skew the roles reverse. The measured overshoot is a change of dominant mode. Relaxed proximal iterations are classical (Pethick et al., 2025); the plateau and the overshoot are observed in the closed loop (Sections 7 and 8).

Schedules converging to an in-window floor are bounded by its exponent (Theorem 6.2).

Theorem 6.2 (Tail rate of floored schedules)**.** *Let $\beta_t \to \beta_{\min} \in (0, 1)$ with $\beta_t \le \bar\beta - \delta$ for some $\delta > 0$ and all large t. Then the asymptotic exponent of $[e; a]_t$ is at most $\rho(M(\beta_{\min}))$, M with the global factor of Lemma 4.3(a). Under the hypotheses of Theorem 6.1 there is $\beta_{\mathrm{s}} > 0$ such that for $\beta_{\min} < \beta_{\mathrm{s}}$ every trajectory not eventually at the equilibrium has tail exponent equal to the modulus of an eigenvalue of the local system at $\beta_{\min}$. Within each mode the trajectories with the fast exponent form a complex subspace of dimension at most one. Outside that subspace on an $\lambda_{\min}$ mode (the slow mode excited) the exponent is the dominant modulus, $1 - \beta_{\min}(1-\rho_J^2) + O(\beta_{\min}^2)$ as $\beta_{\min} \to 0$, and it lies below $\rho(M(\beta_{\min}))$ by at least $\beta_{\min}\rho_J(1-\rho_J)$ to first order, ρ_J the local factor of Lemma 4.3(b).*

The annealing runs of Section 8 report entry times and finite-horizon gaps; neither Kuhn nor Leduc satisfies the interior hypothesis.

Without a floor the same local model settles what a divergent total mass buys (proof in Appendix C).

Theorem 6.3 (Vanishing speeds)**.** *Given the hypotheses of Theorem 6.1, let $\beta_t \in (0, 1)$ be nonincreasing with $\beta_t \to 0$, and call a skew mode* excited *if its component of the state is nonzero. For every $R_0 > 0$ there are $\beta_\circ > 0$ and a neighborhood U of $(z^{\mathrm{NE}}, z^{\mathrm{NE}})$, where z^{NE} is the unique equilibrium, such that if $(z_T, \mu_T) \in U$, $\beta_T \le \beta_\circ$, and on every skew mode the tracking error is at most R_0 times the center error, then (i) on every excited mode the center error stays nonzero and the tracking error is enslaved to it, $e_t/a_t \to 0$ whenever some mode is excited; (ii) if $\sum_t \beta_t = \infty$ and the $\lambda_{\min}$ mode is excited, then $z_t \to Z^\star$ with $-\log \mathrm{dist}(z_t, Z^\star)/\sum_{s<t}\beta_s \to 1 - \rho_J^2$, so no geometric rate exists; (iii) if moreover $\sum_t \beta_t^2 < \infty$, then $\mathrm{dist}(z_t, Z^\star)\, e^{(1-\rho_J^2)\sum_{s<t}\beta_s}$ converges to a positive trajectory-dependent constant; for $\beta_t = c/(1+t)$, $0 < c < 1$, the order is $t^{-c(1-\rho_J^2)}$.*

Constant speeds, floors, and vanishing speeds thus share one local slope, $\beta_t(1-\rho_J^2)$ per step. For the harmonic arm of Section 5 ($c = 0.1$) the exponents are 0.0941 on matching pennies and 0.0923 on rock–paper–scissors. On the evaluation grid these predict least-squares tail slopes of 2.15×10^{-6} and 2.11×10^{-6} per step. The equilibrium distance reproduces them to five digits, and the measured $\mathrm{NashConv}$ slopes of Appendix G agree to within 5% and 0.6%.

7 Rate Degradation

The window guarantee of Section 4 has a structural domain. On Kuhn every arm inside the estimated window converges linearly to an exact equilibrium down to $\mathrm{NashConv} \approx 10^{-10}$. On Leduc, with the same carrier and schedule family, no linear phase is identified. Over 100k–400k the running minimum fits $\mathrm{NashConv} \propto t^{-\alpha}$ with $\alpha \approx 1.8$–2.2. The fitted exponential rate *decreases* in β, and the final $\mathrm{NashConv}$ is non-monotone (Section 8 and Appendix G); these are finite-horizon fits. This section tests one structural lever, $\rho_{\mathrm{in}} \to 1$ through the Lipschitz constant L governing the window (Equation (11)). Kuhn has a non-singleton equilibrium set and still converges linearly, so non-uniqueness alone leaves the rate intact (Section 6). The radial samples of the error-bound modulus are comparable across games; along the Leduc trajectory it is far larger (Appendix G).

Window collapse as $\rho_{\mathrm{in}} \to 1$ The sufficient window Equation (9) is governed by $1-\rho_{\mathrm{in}}$, the small-β expansion of Equation (11). On Leduc, $\rho_{\mathrm{in}} = 0.9485$ at $(\eta, \tau) = (0.2, 0.5)$ with $L = 1.4889$; on Kuhn, $L = 0.661$ gives $\rho_{\mathrm{in}} = 0.917$. The value $\rho_J = 0.9117$ is a plug-in value from a single reference equilibrium (Appendix D); it is not a sample of the modulus γ of Lemma 4.3. With this ρ_J the Leduc window endpoint is $\bar{\beta} = 0.0044$. Even *inside* it, at $\beta = 0.003$, the plug-in rate of $\approx 8.9 \times 10^{-5}$ per step is $3.0\times$ slower than the first-order prediction; Appendix G shows no linear phase beyond $\bar{\beta}$. The plug-in matrix crosses algebraically, as the formula predicts (Table 4): $1-\rho(M(\beta))$ falls from $+8.9\times10^{-5}$ at $\beta = 0.003$ through -6.5×10^{-5} at $\beta = 0.005$ to -2.7×10^{-1} at $\beta = 0.2$. Inside the window the shortfall is second-order: $\beta^2\rho_{\mathrm{in}}/(1-\rho_{\mathrm{in}}) = 1.66\times10^{-4}$ matches the actual shortfall 1.76×10^{-4} below the linear gain 2.65×10^{-4}, so the $3.0\times$ gap is the next term of Equation (11)'s expansion.

The error-bound modulus across the two games. The resolvent factor ρ_J is controlled by the error-bound modulus γ through $\rho_J(\tau) = \gamma\tau/\sqrt{1+\gamma^2\tau^2}$ (Lemma 4.3). Radial samples (Appendix D) give the same order on both games: $\gamma \approx 4.56$ on Leduc against 4.16 on Kuhn (73 and 25 radial points; median ratios 0.62 and 0.95). Sampled maxima bound γ from below. On Leduc the calibrated factor $\rho_J^{\mathrm{eff}} = 0.916$ sits next to the single-point proxy $\rho_J = 0.9117$, and the window barely moves, from $\bar{\beta} = 0.0044$ to $\bar{\beta}^{\mathrm{eff}} = 0.0042$. Every speed that enters $\mathrm{NashConv} < 10^{-2}$ by 50k lies outside both. On Kuhn the calibrated $\rho_J^{\mathrm{eff}} = 0.901$ widens the window from $\bar{\beta} = 0.0057$ to 0.0082 (Appendix G): the linear arms $\beta \le 0.005$ stay inside, the fastest-entering speed $\beta = 0.05$ outside. Kuhn's sampled contraction ratio 0.9325 is itself measured against the whole equilibrium set, and through Equation (7) it implies $\gamma \ge 5.16$ at $\tau = 0.5$. The radial estimate 4.16 and the calibrated factor 0.901 are both below what that sample requires. The two samples measure different quantities and give incompatible calibrations; along the $m = 16$ Leduc run the measured contraction ratio reaches 0.9994 (Appendix G), above every radial calibration.

The inner factor across τ Since ρ_{in} *decreases* in τ while $\rho_J(\tau)$ increases toward 1, the two pull the window in opposite directions on Leduc. Over $\tau \in \{0.5, 1.0, 2.0\}$ the calibrated endpoint shrinks, $\bar{\beta}^{\mathrm{eff}} = 0.0042, 0.0034, 0.0020$, while ρ_{in} drops from 0.9485 to 0.7453 and $\rho_J(\tau)$ climbs from 0.916 to 0.994 (Appendix G). At $\beta = 0.1$ the $\tau = 1.0$ arm is $3\times$ *slower*; the single-point model, which holds $\rho_J = 0.9117$ fixed across τ, widens the window to 0.012 at $\tau = 1.0$ and predicts the opposite ordering. The single-point 2×2 recursion misses the slowdown at both τ; the non-singleton equilibrium-set geometry is one remaining candidate.

Direct intervention on the inner contraction. The window theory puts the tracking budget in ρ_{in} through Proposition 4.4. A direct test holds game, resolvent, initialization, and schedule fixed and runs m inner steps per center update, replacing ρ_{in} by ρ_{in}^m and widening the estimated window to $\bar{\beta}_m = (1-\rho_{\mathrm{in}}^m)(1-\rho_J)/(\rho_{\mathrm{in}}^m(2-\rho_J))$, from 0.0044 to 0.1078 as m grows from 1 to 16 (plug-in windows). On Leduc the widened window predicts a geometric signature by 50k; none appears. At $\beta = 0.04$ the plug-in-window arms $m = 8, 16$ remain non-monotone; at $\beta = 0.1$ $m = 16$ is about $7\times$ below $m = 1$ while every arm stays outside an identifiable geometric regime. The discrepancy

Table 1: Leduc regime identification over 400k iterations (running minimum, $t \geq 100$k); the larger of the two R^2 values is bold.

β	λ_{obs}	α	R^2_{e}	R^2_{p}	NashConv
0.05	9.85×10^{-6}	2.21	**0.964**	0.949	4.25×10^{-5}
0.1	8.12×10^{-6}	1.89	0.905	**0.954**	$\mathbf{3.32 \times 10^{-5}}$
0.2	7.75×10^{-6}	1.82	0.901	**0.974**	5.66×10^{-5}

is large: at $m = 16$, $\beta = 0.04$ the plug-in matrix has $\rho(M_{16}) = 0.9977$, whose 50k-step bound decays by a factor of about e^{114}, while the measured tail oscillates near 10^{-3}. The plug-in bound is therefore not applicable as stated. Instrumenting this run (Appendix G) locates the failure in the center row. The measured contraction ratio of J toward $Z^\star$ rises from 0.906 to 0.9994 along the run, against the plug-in $\rho_J = 0.9117$; with it, $\rho(M_{16}) = 1.0012$, so the comparison bound certifies no decay while the dynamics decay at a per-block ratio of 0.99995. The inner contraction factor shapes the transient; the break persists at every m tested.

Where ρ_{in} is bounded away from 1 (Kuhn), convergence is linear inside the estimated window and, empirically, outside it. Where ρ_{in} approaches 1 (Leduc), every single-step ($m = 1$) speed that progresses within the horizon lies outside the estimated window, and the finite-horizon behavior is power-law-like on the non-singleton $Z^\star$. The windows here are plug-in estimates; within the tested ranges neither τ nor the inner contraction produced an identifiable geometric regime.

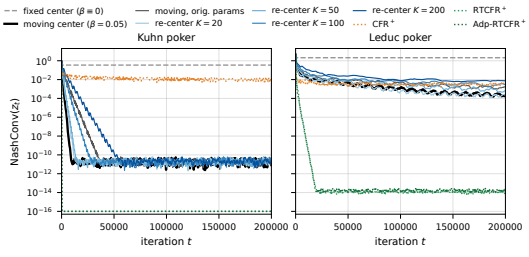
8 Experiments

Four families test the theory: a constant-β scan on Kuhn (window and slope), power-law fits on Leduc (degradation), annealing onto floors, and a four-game baseline. We evaluate the last policy under the unnormalized saddle gap NashConv (Equation (2)); *measured entry* is the first iterate with $\text{NashConv} < 10^{-6}$.

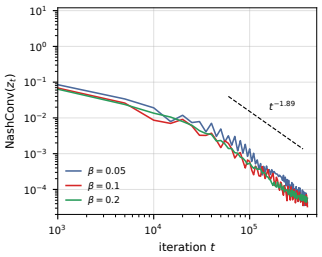
Constant speeds on Kuhn: the window and the slope Ten constant schedules $\beta \in \{0.001, \ldots, 0.2\}$ run for 30k iterations at $(\eta, \tau) = (0.2, 0.5)$; the estimated window is $\bar{\beta} = 0.0057$ with $\rho_{\text{in}} = 0.917$ and $\rho_J = 0.9325$ (Theorem 4.5 and Section 7). All four arms inside the estimated window decay linearly with no floor, $\beta = 0.003$ to 2.3×10^{-7} and $\beta = 0.005$ to 1.1×10^{-10}. Their slopes grow linearly in β at ratio ≈ 0.14–0.17. This ratio is close to the closed-form $1 - \rho_J^2 = 0.130$ of Theorem 6.1 and twice the worst-case $1 - \rho_J = 0.068$ of Equation (11). Kuhn's equilibrium is on the boundary, outside Theorem 6.1's interiority hypothesis, so the agreement is empirical. The six arms beyond the window also converge, with the same tail slope and earlier entry (Appendix G).

Leduc: power-law fits and long horizons A ten-arm scan over the same grid (50k) inverts the window ordering. No arm with $\beta \leq 0.008$ reaches $\text{NashConv} < 10^{-2}$; larger β enters earlier, $\beta = 0.01$ at iteration 45,900 and $\beta = 0.1$ at 9,800, outside the estimated window. Table 1 fits the 100k–400k running minimum (Figure 1b), reporting the fitted exponential rate λ_{obs} per iteration, the fitted power-law exponent α, and the coefficients of determination R^2_{e} and R^2_{p} of the two fits. The power-law fit wins in R^2_{p} for $\beta \in \{0.1, 0.2\}$ and the exponential in R^2_{e} at $\beta = 0.05$. The fitted exponential rate λ_{obs} decreases in β, and the exponent α shrinks with the schedule, $2.21 \to 1.82$. The best final NashConv ($\beta = 0.1$) is the crossing of the two. At 500k the ordering reverses: the moving center $\beta = 0.05$ reaches 2.12×10^{-5}, overtaking every re-centering arm (4.32×10^{-5} at $K = 20$, 6.65×10^{-5} at $K = 50$). The degradation is not a reference artifact: the iterate approaches a point of $Z^\star$ other than the one the single-point proxy is built on (Appendix G).

Nearest-neighbor baselines across four games. In Table 2, layer A instantiates Equation (4) on the common carrier: fixed center $\beta \equiv 0$ (Sokota et al., 2023); moving center $\beta = 0.05$, and at the original MMD parameters $(0.1, 1.0)$; and periodic re-centering $\beta_t = \mathbf{1}[(t+1) \bmod K = 0]$, $K \in \{20, 50, 100, 200\}$ (APMD-style (Abe et al., 2024)). Layer B runs our reimplementations of these algorithms with native regret-matching inner loops (CFR$^+$ (Tammelin, 2014), RTCFR$^+$ (Meng et al., 2025), Adp-RTCFR$^+$ (Ren et al., 2025)), RTCFR$^+$ in the form of Algorithm 1 of Ren et al. (2025) at its published hyperparameters and Adp-RTCFR$^+$ at the RTCFR$^+$ values; the propagating



(a) Trajectories behind Table 2 (200k; layer A solid, layer B dotted).



(b) Leduc over 400k, log–log, against $t^{-1.89}$.

Figure 1: Window behavior on Kuhn and power-law degradation on Leduc.

Table 2: Final NashConv (entry iteration into NashConv $< 10^{-6}$; iteration counts, not wall-clock time); horizons 200k on Kuhn/Leduc, 100k and 50k on 4- and 5-sided Liar's Dice; 10^{-16} and 10^{-14} are floating-point floors, reproduced to 9×10^{-16} by OpenSpiel on the saved policies (Appendix D); dashes mean no entry; $K = 200$ on 5-sided Liar's Dice enters at 53,600.

	Method	Kuhn	Leduc	Liar's Dice 4-sided	Liar's Dice 5-sided
A	fixed center	3.5×10^{-1} (–)	2.10 (–)	1.41 (–)	1.58 (–)
	moving center	8.4×10^{-12} (5200)	2.30×10^{-4} (–)	10^{-16} (4600)	10^{-16} (8000)
	orig. params	1.5×10^{-11} (18800)	1.69×10^{-3} (–)	10^{-16} (9000)	10^{-16} (16400)
	re-center K=20	1.5×10^{-11} (6400)	2.12×10^{-4} (–)	10^{-16} (3600)	10^{-16} (6400)
	re-center K=50	2.0×10^{-11} (8000)	6.42×10^{-4} (–)	10^{-16} (7600)	10^{-16} (13500)
	re-center K=100	2.6×10^{-11} (14400)	2.95×10^{-3} (–)	10^{-16} (15000)	10^{-16} (26800)
	re-center K=200	2.8×10^{-11} (28800)	7.44×10^{-3} (–)	10^{-16} (29800)	9.97×10^{-5} (–)
B	CFR$^+$	5.9×10^{-3} (–)	3.4×10^{-3} (–)	10^{-16} (200)	10^{-16} (100)
	RTCFR$^+$	10^{-16} (400)	1.7×10^{-14} (6000)	10^{-16} (200)	10^{-16} (300)
	Adp-RTCFR$^+$	10^{-16} (400)	1.7×10^{-14} (6000)	10^{-16} (200)	10^{-16} (100)

variant of RTCFR$^+$ enters on the same grid at the iterations listed in Appendix E. The layers share the metric, differ in geometry, and are never pooled.

Independent validation. The game construction and the gap evaluator are checked against independent computations (Appendix D). The sequence-form linear program returns the game values, $-1/18$ on Kuhn and -0.085606 on Leduc against OpenSpiel's -0.085603. A synchronous tree traversal matches each of our games to an OpenSpiel game node by node, asserting the acting player, the action counts, the chance probabilities, and the terminal payoffs, and recording the information-set bijection without parsing strings. Every evaluated policy of every layer-B entry agrees between the two evaluators to 8.9×10^{-16}.

9 Conclusion

The reference strategy is the knob every last-iterate method in this family turns, and the coordinate on which the family becomes analyzable: one schedule underlies the fixed magnet of MMD, the periodic re-centering of APMD, and the hard resets of RTCFR$^+$, and the theorems hold for all of them at once. The interior case, where earlier analyses stop at asymptotics, gains an explicit rate: constant speeds inside Equation (9) converge linearly in the last iterate to an *exact* Nash equilibrium of the original game, and the endpoint is sharp. Outside the window, summable schedules freeze at the computable resolvent of the center's limit with a positive floor Equation (17), and every schedule shares the slope $\beta_t(1-\rho_J^2)$, a closed form in the speed, the proximal weight, and the skew spectrum that reproduces measured slopes to three significant figures with no fitted constant. Across four games, ten methods, and twenty-eight carrier-matched entries the moving center reaches a saddle gap below 10^{-6} wherever any other common-carrier method does.

AI USE STATEMENT

In this work, we used generative AI tools to polish the prose and to write and run the experiment code. We did not use them to generate the research ideas or to design the methodology, and no other disclosure requirement applies. We reviewed all AI-assisted work against the underlying theorems, numbers, and figures, and verified the code by rerunning the experiments. We take responsibility for the final content.

REPRODUCIBILITY STATEMENT

Theoretical claims are proved in the appendix: the window lemmas and Theorem 4.5 in Appendix A, the freezing theorem in Appendix B, and the local-rate and tail theorems in Appendix C. The experimental protocol is in Appendix D, with Layer-B reproduction and the deferred studies in Appendices E and G. Rate predictions use no fitted constant, and every fitted quantity is labeled as a fit. An anonymized code package ships the drivers and the recorded reference tables behind the reported numbers and the figure series.

REFERENCES

Kenshi Abe, Kaito Ariu, Mitsuki Sakamoto, and Atsushi Iwasaki. Adaptively perturbed mirror descent for learning in games. In *Proceedings of the 41st International Conference on Machine Learning*, pp. 31–80, 2024.

Kenshi Abe, Mitsuki Sakamoto, Kaito Ariu, and Atsushi Iwasaki. Asymmetric perturbation in solving bilinear saddle-point optimization. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=fyae4jQ9Lw.

Jacob Abernethy, Peter L Bartlett, and Elad Hazan. Blackwell approachability and no-regret learning are equivalent. In *Proceedings of the 24th Annual Conference on Learning Theory*, pp. 27–46. JMLR Workshop and Conference Proceedings, 2011.

James Alcala, Yat Tin Chow, and Mahesh Sunkula. Stochastic moving anchor algorithms and a popov's scheme with moving anchor, 2025a. URL https://arxiv.org/abs/2506.07290.

James K. Alcala, Yat Tin Chow, and Mahesh Sunkula. Moving anchor extragradient methods for smooth structured minimax problems, 2025b. URL https://arxiv.org/abs/2308.12359.

Kimon Antonakopoulos, Veronica Belmega, and Panayotis Mertikopoulos. Adaptive extra-gradient methods for min-max optimization and games. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=R0a0kFI3dJx.

Heinz H Bauschke and Patrick L Combettes. Correction to: convex analysis and monotone operator theory in hilbert spaces. In *Convex analysis and monotone operator theory in Hilbert spaces*, pp. C1–C4. Springer, 2020.

Michael Bowling, Neil Burch, Michael Johanson, and Oskari Tammelin. Heads-up limit hold'em poker is solved. *Science*, 347(6218):145–149, 2015.

Mario Bravo, Roberto Cominetti, and Jongmin Lee. Minimax-optimal halpern iterations for lipschitz maps, 2026. URL https://arxiv.org/abs/2601.15996.

George W Brown. Iterative solution of games by fictitious play. *Act. Anal. Prod Allocation*, 13(1): 374, 1951.

Noam Brown and Tuomas Sandholm. Solving imperfect-information games via discounted regret minimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pp. 1829–1836, 2019.

Yang Cai and Weiqiang Zheng. Last-iterate convergence of anchored gradient descent, 2026. URL https://arxiv.org/abs/2604.12235.

Yang Cai, Gabriele Farina, Julien Grand-Clément, Christian Kroer, Chung-Wei Lee, Haipeng Luo, and Weiqiang Zheng. Fast last-iterate convergence of learning in games requires forgetful algorithms. *Advances in Neural Information Processing Systems*, 37:23406–23434, 2024.

Yang Cai, Haipeng Luo, Chen-Yu Wei, and Weiqiang Zheng. From average-iterate to last-iterate convergence in games: A reduction and its applications. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=mG9xItYI0D.

Yang Cai, Constantinos Daskalakis, Haipeng Luo, Chen-Yu Wei, and Weiqiang Zheng. Proximal regret and proximal correlated equilibria: A new tractable solution concept for online learning and games. In *Proceedings of the 58th Annual ACM Symposium on Theory of Computing*, pp. 1639–1650, 2026.

Shicong Cen, Yuting Wei, and Yuejie Chi. Fast policy extragradient methods for competitive games with entropy regularization. *Advances in Neural Information Processing Systems*, 34:27952–27964, 2021.

Darshan Chakrabarti, Julien Grand-Clément, and Christian Kroer. Extensive-form game solving via blackwell approachability on treeplexes. *Advances in Neural Information Processing Systems*, 37:35257–35287, 2024.

Zihao Chen. A unifying view of anchoring via operator-side tikhonov regularization, 2026. URL https://arxiv.org/abs/2605.30905.

C Daskalakis and Ioannis Panageas. Last-iterate convergence: Zero-sum games and constrained min-max optimization. In *10th Innovations in Theoretical Computer Science (ITCS) conference, ITCS 2019*, 2019.

Gabriele Farina, Christian Kroer, Noam Brown, and Tuomas Sandholm. Stable-predictive optimistic counterfactual regret minimization. In *International conference on machine learning*, pp. 1853–1862. PMLR, 2019.

Gabriele Farina, Julien Grand-Clément, Christian Kroer, Chung-Wei Lee, and Haipeng Luo. Regret matching+:(in) stability and fast convergence in games. *Advances in Neural Information Processing Systems*, 36:61546–61572, 2023.

Eduard Gorbunov, Adrien Taylor, and Gauthier Gidel. Last-iterate convergence of optimistic gradient method for monotone variational inequalities. *Advances in neural information processing systems*, 35:21858–21870, 2022.

Ke Guo and Deren Han. Nonsymmetric proximal point algorithm with moving proximal centers for variational inequalities: Convergence analysis. *Applied Numerical Mathematics*, 147:1–18, 2020.

James Hannan. Approximation to bayes risk in repeated play. *Contributions to the Theory of Games*, 3(2):97–139, 1957.

Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

Marc Lanctot, Kevin Waugh, Martin Zinkevich, and Michael Bowling. Monte carlo sampling for regret minimization in extensive games. *Advances in neural information processing systems*, 22, 2009.

Marc Lanctot, Edward Lockhart, Jean-Baptiste Lespiau, Vinicius Zambaldi, Satyaki Upadhyay, Julien Pérolat, Sriram Srinivasan, Finbarr Timbers, Karl Tuyls, Shayegan Omidshafiei, Daniel Hennes, Dustin Morrill, Paul Muller, Timo Ewalds, Ryan Faulkner, János Kramár, Bart De Vylder, Brennan Saeta, James Bradbury, David Ding, Sebastian Borgeaud, Matthew Lai, Julian Schrittwieser, Thomas Anthony, Edward Hughes, Ivo Danihelka, and Jonah Ryan-Davis. Openspiel: A framework for reinforcement learning in games. *arXiv preprint arXiv:1908.09453*, 2019.

Luis Leal. Which nash equilibrium? solver-dependent selection on zero-sum nash polytopes, 2026. URL https://arxiv.org/abs/2606.28308.

Chung-Wei Lee, Christian Kroer, and Haipeng Luo. Last-iterate convergence in extensive-form games. *Advances in Neural Information Processing Systems*, 34:14293–14305, 2021.

Mingyang Liu, Asuman E Ozdaglar, Tiancheng Yu, and Kaiqing Zhang. The power of regularization in solving extensive-form games. In *The Eleventh International Conference on Learning Representations*, 2023.

Weiming Liu, Huacong Jiang, Bin Li, and Houqiang Li. Equivalence analysis between counterfactual regret minimization and online mirror descent. In *International Conference on Machine Learning*, pp. 13717–13745. PMLR, 2022.

Yixin Liu, Argyris Oikonomou, Weiqiang Zheng, Yang Cai, and Arman Cohan. COMAL: A convergent meta-algorithm for aligning LLMs with general preferences. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=OsrE5DJ9Fu.

Wataru Masaka, Mitsuki Sakamoto, Kenshi Abe, Kaito Ariu, Tuomas Sandholm, and Atsushi Iwasaki. On the power of perturbation under sampling in solving extensive-form games. *arXiv preprint arXiv:2501.16600*, 2025.

Linjian Meng, Tianpei Yang, Youzhi Zhang, Zhenxing Ge, Shangdong Yang, Tianyu Ding, Wenbin Li, Bo An, and Yang Gao. Efficient last-iterate convergence in solving extensive-form games. In *Advances in Neural Information Processing Systems*, 2025. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/918c67cd7910578f1637beb56717f455-Abstract-Conference.html.

Julien Perolat, Bart De Vylder, Daniel Hennes, Eugene Tarassov, Florian Strub, Vincent de Boer, Paul Muller, Jerome T Connor, Neil Burch, Thomas Anthony, et al. Mastering the game of stratego with model-free multiagent reinforcement learning. *Science*, 378(6623):990–996, 2022.

Thomas Pethick, Ioannis Mavrothalassitis, and Volkan Cevher. Efficient interpolation between extragradient and proximal methods for weak mvis. In *The Thirteenth International Conference on Learning Representations*, 2025.

Mihály Pituk. More on Poincaré's and Perron's theorems for difference equations. *The Journal of Difference Equations and Applications*, 8(3):201–216, 2002.

Hang Ren, Yulin Wu, Shuhan Qi, Jiajia Zhang, Xiaozhen Sun, Tianzi Ma, and Xuan Wang. Efficient last-iterate convergence in regret minimization via adaptive reward transformation. *arXiv preprint arXiv:2509.13653*, 2025.

Stephen M. Robinson. Some continuity properties of polyhedral multifunctions. In *Mathematical Programming at Oberwolfach*, volume 14 of *Mathematical Programming Studies*, pp. 206–214. Springer, 1981. doi: 10.1007/BFb0120929.

Can Savcı. Garip: A running-average moving reference for last-iterate self-play in two-player zero-sum games, 2026. URL https://arxiv.org/abs/2606.22688.

Samuel Sokota, Ryan D'Orazio, J Zico Kolter, Nicolas Loizou, Marc Lanctot, Ioannis Mitliagkas, Noam Brown, and Christian Kroer. A unified approach to reinforcement learning, quantal response equilibria, and two-player zero-sum games. In *The Eleventh International Conference on Learning Representations*, 2023.

Anja Surina, Arun Suggala, George Tsoukalas, Anton Kovsharov, Sergey Shirobokov, Francisco J. R. Ruiz, Pushmeet Kohli, and Swarat Chaudhuri. An improved last-iterate convergence rate for anchored gradient descent ascent, 2026. URL https://arxiv.org/abs/2604.03782.

Oskari Tammelin. Solving large imperfect information games using cfr+. *arXiv preprint arXiv:1407.5042*, 2014.

Bernhard Von Stengel. Efficient computation of behavior strategies. *Games and Economic Behavior*, 14(2):220–246, 1996.

Mingzhi Wang, Chengdong Ma, Qizhi Chen, Linjian Meng, Yang Han, Jiancong Xiao, Zhaowei Zhang, Jing Huo, Weijie J Su, and Yaodong Yang. Magnetic preference optimization: Achieving last-iterate convergence for language model alignment. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=PDnEDS244P.

Chen-Yu Wei, Chung-Wei Lee, Mengxiao Zhang, and Haipeng Luo. Linear last-iterate convergence in constrained saddle-point optimization. In *International Conference on Learning Representations*, 2021.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th international conference on machine learning (icml-03)*, pp. 928–936, 2003.

Martin Zinkevich, Michael Johanson, Michael Bowling, and Carmelo Piccione. Regret minimization in games with incomplete information. *Advances in neural information processing systems*, 20, 2007.

A Proofs for Section 4

Throughout the appendix, $\|\cdot\|$ is the Euclidean norm, Π_Z the Euclidean projection onto the joint treeplex $Z = X \times Y$, and $T \coloneqq F + N_Z$. We use repeatedly that Π_Z is nonexpansive and that F is linear and skew (Section 3).

Proof of Lemma 4.1 (inner contraction)

Proof. Fix μ and let $\delta = z - z'$. By linearity of F and skewness, $\langle \delta, F\delta \rangle = 0$, so

$$\|\delta - \eta F\delta\|^2 = \|\delta\|^2 - 2\eta\langle \delta, F\delta\rangle + \eta^2\|F\delta\|^2 = \|\delta\|^2 + \eta^2\|F\delta\|^2 \le (1 + \eta^2 L^2)\,\|\delta\|^2.$$

Since the center term $\eta\tau\mu$ cancels in the difference and Π_Z is nonexpansive,

$$\|G_\mu(z) - G_\mu(z')\| \;\le\; \frac{\|\delta - \eta F\delta\|}{1+\eta\tau} \;\le\; \frac{\sqrt{1+\eta^2L^2}}{1+\eta\tau}\,\|\delta\| \;=\; \rho_{\mathrm{in}}\|\delta\|.$$

The equivalence in Equation (6) is algebra: $1 + \eta^2L^2 < (1+\eta\tau)^2 \iff \eta L^2 < 2\tau + \eta\tau^2 \iff \eta(L^2 - \tau^2) < 2\tau$. Under it, G_μ is a contraction on the complete set Z, and the Banach fixed-point theorem gives the unique fixed point $z^\star_\tau(\mu)$ and the geometric estimate. □

The cross-term cancellation is exact, but $\|F\delta\| = L\|\delta\|$ is attained along the top singular directions of A. On matching pennies and rock–paper–scissors these directions span the simplex tangent space, so the bound is attained: the measured worst-case contraction ratios 0.979121 and 0.962091 equal ρ_{in} to six digits. On the Kuhn treeplex the measured ratio 0.909640 sits below $\rho_{\mathrm{in}} = 0.917$, so the bound is conservative there; theorem parameters and experiment parameters are stated separately in Section 8.

Proof of Lemma 4.2 (the fixed point is a resolvent)

Proof. Let $z^\star$ be a fixed point of G_μ and $w = \big(z^\star + \eta\tau\mu - \eta F(z^\star)\big)/(1+\eta\tau)$, so $z^\star = \Pi_Z(w)$. The projection characterization $\langle z - z^\star, w - z^\star\rangle \le 0$ for all $z \in Z$, together with

$$w - z^\star = \frac{z^\star + \eta\tau\mu - \eta F(z^\star) - (1+\eta\tau)z^\star}{1+\eta\tau} = -\,\frac{\eta}{1+\eta\tau}\big(F(z^\star) + \tau(z^\star - \mu)\big),$$

is equivalent to $\langle z - z^\star,\, F(z^\star) + \tau(z^\star - \mu)\rangle \ge 0$ for all $z \in Z$, i.e. $-\big(F(z^\star) + \tau(z^\star - \mu)\big) \in N_Z(z^\star)$, i.e.

$$0 \;\in\; F(z^\star) + N_Z(z^\star) + \tau(z^\star - \mu) \quad\iff\quad \mu \;\in\; z^\star + \tau^{-1}T(z^\star) \quad\iff\quad z^\star = J(\mu).$$

F is monotone (skew) and continuous, N_Z is maximally monotone, and $\operatorname{dom} F = \mathbb{R}^d$, so T is maximally monotone; by Minty's theorem (Bauschke & Combettes, 2020) its resolvent is single-valued on the whole space and firmly nonexpansive. $\mathrm{Fix}(J) = Z^\star$ because $J(\mu) = \mu \iff 0 \in T(\mu)$. Firm nonexpansiveness implies nonexpansiveness of $\mu \mapsto z^\star_\tau(\mu)$. □

Proof of Lemma 4.3 (resolvent contraction)

Proof of part (a). Write $z^+ = J(\mu)$, $r = \|\mu - z^+\|$, $a = \operatorname{dist}(\mu, Z^\star)$, $d = \operatorname{dist}(z^+, Z^\star)$. Two independent bounds hold.

Error-bound leg. By the resolvent inclusion, $\tau(\mu - z^+) \in T(z^+)$, so $\operatorname{dist}(0, T(z^+)) \le \tau r$ and metric subregularity gives $d \le \gamma\tau r$.

Firm-nonexpansiveness leg. Let $\nu = \Pi_{Z^\star}(\mu)$, so $J(\nu) = \nu$ and $(I - J)\nu = 0$. Firm nonexpansiveness of J gives

$$\|z^+ - \nu\|^2 \;\le\; \|\mu - \nu\|^2 - \|(I - J)\mu - (I - J)\nu\|^2 \;=\; a^2 - r^2,$$

hence $d \le \sqrt{a^2 - r^2}$.

Combining, $d \le \min\bigl(\gamma\tau r, \sqrt{a^2 - r^2}\bigr)$ for the realized r; the bound is worst over r where the two legs meet, at $r^2 = a^2/(1 + \gamma^2\tau^2)$, giving $d \le \gamma\tau a/\sqrt{1 + \gamma^2\tau^2} = \rho_J\, a$ with $\rho_J < 1$.

For the hypothesis: F is linear and Z is a polytope, so the graph of T is a finite union of polyhedral sets and T is a polyhedral multifunction; metric subregularity then holds at every point of $Z^\star$ (Robinson, 1981). Each point of $Z^\star$ admits a local modulus and a neighborhood in Z; since $Z^\star$ is compact (closed in the compact Z), finitely many such neighborhoods cover $Z^\star$; call their union U. On the compact remainder $Z \setminus U$, when it is nonempty, T is closed and $T^{-1}(0) = Z^\star \subseteq U$, so by closedness of the graph of T and compactness the infimum of $\operatorname{dist}(0, T(z))$ there is attained and positive: $\operatorname{dist}(0, T(z)) \ge c > 0$, while $\operatorname{dist}(z, Z^\star) \le \operatorname{diam}(Z)$. Taking γ as the maximum of the finitely many local moduli and, when the remainder is nonempty, $\operatorname{diam}(Z)/c$ gives $\operatorname{dist}(z, Z^\star) \le \gamma \operatorname{dist}(0, T(z))$ on all of Z, i.e. a modulus valid along the entire trajectory. □

Proof of part (b). Suppose $Z^\star = \{z_{\mathrm{NE}}\}$ with z_{NE} in the relative interior of Z. In a neighborhood of z_{NE} the active inequality constraints are empty, so Z agrees locally with its affine hull and Π_Z acts as the affine projection onto that affine subspace, a linear map once coordinates are centered at a point of it. In tangent-space coordinates centered at z_{NE}, let V be the tangent space and P_V the orthogonal projection onto it; writing $F_0 = P_V F P_V$ for the tangential compression of F (skew on V, with $P_V F(z_{\mathrm{NE}}) = 0$ by the interior equilibrium condition), the resolvent inclusion $0 \in F(z^\star) + \tau(z^\star - \mu) + N_Z(z^\star)$ (whose normal cone is $V^\perp$ at the interior point), projected onto V, gives $0 = F_0 z^\star + \tau(z^\star - \mu)$, which solves to

$$z^\star - z_{\mathrm{NE}} \;=\; \tau\,(\tau I + F_0)^{-1}(\mu - z_{\mathrm{NE}}).$$

F_0 is skew, so $\tau I + F_0$ is normal with eigenvalues $\tau \pm i\lambda_k$, where $\lambda_k > 0$ are the skew strengths; uniqueness of the equilibrium forces F_0 to have no zero eigenvalue on V: if $F_0 v = 0$ for a nonzero tangent direction v, then $z_{\mathrm{NE}} + \varepsilon v$ stays relatively interior for small ε, and since an interior point of Z is an equilibrium exactly when $P_V F$ vanishes there, $P_V F(z_{\mathrm{NE}} + \varepsilon v) = \varepsilon F_0 v = 0$ makes it a second equilibrium. Hence

$$\|\tau(\tau I + F_0)^{-1}\| \;=\; \frac{\tau}{\min_k |\tau + i\lambda_k|} \;=\; \frac{\tau}{\sqrt{\tau^2 + \lambda_{\min}^2}},$$

which is the claimed local ρ_J (the minimum runs over the nonzero skew strengths; a trivial tangent space gives the constant resolvent and $\rho_J = 0$); for $V \neq \{0\}$, comparison with part (a) gives the local modulus $\gamma = 1/\lambda_{\min}$. □

Proof of Proposition 4.4 (coupled recursion)

Proof. Fix t and write $\beta = \beta_t$, $z^\star(\mu) = z^\star_\tau(\mu) = J(\mu)$, $\pi = \Pi_{Z^\star}(\mu_t)$.

Preliminary bound. Since J is firmly nonexpansive, so is $I - J$; with $(I - J)\pi = 0$,

$$\|z^\star(\mu_t) - \mu_t\| = \|(I - J)\mu_t - (I - J)\pi\| \le \|\mu_t - \pi\| = a_t. \tag{14}$$

Inner step. With μ_t frozen, Lemma 4.1 gives $\|z_{t+1} - z^\star(\mu_t)\| \le \rho_{\mathrm{in}} e_t$.

Center row. μ_{t+1} is a convex combination of μ_t and z_{t+1}, and $\mathrm{dist}(\cdot, Z^\star)$ is convex because $Z^\star$ is convex, so

$$\begin{aligned} a_{t+1} &\le (1-\beta)\, a_t + \beta\, \mathrm{dist}(z_{t+1}, Z^\star) \\ &\le (1-\beta)\, a_t + \beta\big[\|z_{t+1} - z^\star(\mu_t)\| + \mathrm{dist}(z^\star(\mu_t), Z^\star)\big] \\ &\le (1-\beta)\, a_t + \beta\big[\rho_{\mathrm{in}} e_t + \rho_J a_t\big] \;=\; \big(1 - \beta(1-\rho_J)\big) a_t + \beta\rho_{\mathrm{in}} e_t, \end{aligned}$$

using Lemma 4.3 in the last inequality.

Tracking row. By the triangle inequality and nonexpansiveness of $\mu \mapsto z^\star(\mu)$ (Lemma 4.2),

$$e_{t+1} \le \|z_{t+1} - z^\star(\mu_t)\| + \|z^\star(\mu_t) - z^\star(\mu_{t+1})\| \le \rho_{\mathrm{in}} e_t + \|\mu_{t+1} - \mu_t\| = \rho_{\mathrm{in}} e_t + \beta\|z_{t+1} - \mu_t\|.$$

The last factor splits as $\|z_{t+1} - \mu_t\| \le \|z_{t+1} - z^\star(\mu_t)\| + \|z^\star(\mu_t) - \mu_t\| \le \rho_{\mathrm{in}} e_t + a_t$ by Equation (14), giving $e_{t+1} \le \rho_{\mathrm{in}}(1+\beta)\, e_t + \beta a_t$.

Bridge. $\mathrm{dist}(z_t, Z^\star) \le e_t + \mathrm{dist}(z^\star(\mu_t), Z^\star) \le e_t + \rho_J a_t$ by Lemma 4.3, and $\mathrm{NashConv}$ vanishes on $Z^\star$ and is L_{gap}-Lipschitz (Section 3), so $\mathrm{NashConv}(z_t) \le L_{\mathrm{gap}}(e_t + \rho_J a_t)$. □

The bridge constant is $L_{\mathrm{gap}} \le \sqrt{2}\, L\, R$: writing $\mathrm{NashConv}(z) = \phi(y) + \psi(x)$ with $\phi(y) = \max_{x' \in X} x'^\top A y$ and $\psi(x) = \max_{y' \in Y}(-x^\top A y')$, each is a max of linear forms, so $|\phi(y_1) - \phi(y_2)| \le R_X L \|y_1 - y_2\|$ and $|\psi(x_1) - \psi(x_2)| \le R_Y L \|x_1 - x_2\|$, where $R_X = \max_{x \in X}\|x\|$ and $R_Y = \max_{y \in Y}\|y\|$; Cauchy–Schwarz on the two increments gives $L_{\mathrm{gap}} \le L\sqrt{R_X^2 + R_Y^2}$, and bounding both radii by $R = \max(R_X, R_Y)$ gives the stated $L_{\mathrm{gap}} \le \sqrt{2}\, L\, R$, with $R = 1$ for simplex games.

Numerically, the two inequalities of the recursion are checked step by step, with zero violations, over three matrix games (unique equilibria) at $\beta \in \{0.01, 0.05, 0.2\}$ and on the Kuhn treeplex at $\beta \in \{0.002, 0.01, 0.05\}$, 399 steps each; a_t is the true distance to $Z^\star = X^\star \times Y^\star$ (by quadratic-programming projection) and ρ_J the sampled value, so the comparison is checked numerically at every step.

Proof of Theorem 4.5 (windowed linear convergence) and of the expansion Equation (11)

Proof of Theorem 4.5. Write $M(\beta) = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right]$ with $p = \rho_{\mathrm{in}}(1+\beta)$, $q = \beta$, $r = \beta\rho_{\mathrm{in}}$, $s = 1 - \beta(1-\rho_J)$, all nonnegative.

Window $\iff$ spectral radius below one. The eigenvalues of the nonnegative 2×2 matrix $M(\beta)$ are real, $\lambda_\pm = \frac{p+s}{2} \pm \sqrt{\left(\frac{p-s}{2}\right)^2 + qr}$, and $\rho(M(\beta)) = \lambda_+$. The characteristic polynomial $\chi(\lambda) = (\lambda - p)(\lambda - s) - qr$ satisfies $\chi(1) = (1-p)(1-s) - qr$. If $p < 1$, $s < 1$, and $(1-p)(1-s) > qr$, then $\chi(1) > 0$ and $\frac{p+s}{2} < 1$, so both roots lie below one; conversely $\lambda_+ < 1$ forces $\chi(1) > 0$ and $p, s \le \lambda_+ < 1$. Substituting,

$$\begin{aligned} (1-p)(1-s) > qr \;&\iff\; \big(1 - \rho_{\mathrm{in}} - \beta\rho_{\mathrm{in}}\big)\,\beta(1-\rho_J) > \beta^2\rho_{\mathrm{in}} \\ &\iff\; (1-\rho_{\mathrm{in}})(1-\rho_J) > \beta\rho_{\mathrm{in}}(2-\rho_J) \;\iff\; \beta < \bar\beta, \end{aligned}$$

where the middle step divides by $\beta > 0$ and rearranges $\beta\rho_{\mathrm{in}}(1-\rho_J) + \beta\rho_{\mathrm{in}} = \beta\rho_{\mathrm{in}}(2-\rho_J)$. The condition $\beta < \bar\beta$ also forces $p < 1$: $\bar\beta < (1-\rho_{\mathrm{in}})/\rho_{\mathrm{in}}$, so $\beta\rho_{\mathrm{in}} < 1 - \rho_{\mathrm{in}}$; and $s < 1$ holds for every $\beta > 0$ since $\rho_J < 1$.

Rate. Iterating Equation (8) elementwise (all entries nonnegative) gives $[e; a]_t \le M(\beta)^t [e; a]_0$, hence $\|[e; a]_t\| \le \|M(\beta)^t\|\, \|[e_0; a_0]\|$. The constant is explicit: with $D = \mathrm{diag}(1, \sqrt{\rho_{\mathrm{in}}})$ the similar matrix $D^{-1}M(\beta)D$ has both off-diagonal entries equal to $\beta\sqrt{\rho_{\mathrm{in}}}$ ($q\sqrt{\rho_{\mathrm{in}}}$ and $r/\sqrt{\rho_{\mathrm{in}}}$), so it is symmetric, shares the eigenvalues of $M(\beta)$, and has Euclidean operator norm $\rho(M(\beta))$; therefore

$$\|M(\beta)^t\| = \big\|D\,(D^{-1}M(\beta)D)^t D^{-1}\big\| \le \|D\|\,\|D^{-1}\|\,\rho(M(\beta))^t = \rho_{\mathrm{in}}^{-1/2}\,\rho(M(\beta))^t$$

for every $t \ge 0$. By the bridge in Proposition 4.4 and Cauchy–Schwarz,

$$\begin{aligned} \mathrm{NashConv}(z_t) &\le L_{\mathrm{gap}}(e_t + \rho_J a_t) \le L_{\mathrm{gap}}\sqrt{1+\rho_J^2}\,\big\|[e; a]_t\big\| \\ &\le L_{\mathrm{gap}}\sqrt{1+\rho_J^2}\;\rho_{\mathrm{in}}^{-1/2}\,\rho(M(\beta))^t\,\big\|[e_0; a_0]\big\|. \end{aligned}$$

Point convergence. To see that z_t converges to a *specific* point of $Z^\star$, use firm nonexpansiveness of J: for $\nu_t = \Pi_{Z^\star}(\mu_t)$, $\|(I-J)(\mu_t)\|^2 \le \|\mu_t-\nu_t\|^2 - \|J(\mu_t)-\nu_t\|^2 \le a_t^2$, so $\|z_t-\mu_t\| \le e_t+a_t$, and applying the same bound at time $t+1$ gives $\|z_{t+1}-\mu_{t+1}\| \le e_{t+1}+\|(I-J)\mu_{t+1}\| \le e_{t+1}+a_{t+1}$.

$$\|z_{t+1}-z_t\| \;\le\; \|z_{t+1}-\mu_{t+1}\|+\|\mu_{t+1}-\mu_t\|+\|\mu_t-z_t\| \;\le\; (e_{t+1}+a_{t+1})+\beta(e_t+a_t)+(e_t+a_t),$$

with both $e.$ and $a.$ geometrically summable, z_t is Cauchy; since $\operatorname{dist}(z_t, Z^\star) \to 0$ with $Z^\star$ closed, $z_t \to z_\infty \in Z^\star$.

Endpoint sharpness. When $\bar\beta \le 1$, so that $\bar\beta$ lies in the schedule domain and $s = 1-\bar\beta(1-\rho_J) \ge \rho_J \ge 0$, the derivation above at $\beta = \bar\beta$ gives $(1-p)(1-s) = qr$, hence $\chi(1) = 0$: $\lambda = 1$ is an eigenvalue of the nonnegative matrix $M(\beta)$, and the other eigenvalue $p+s-1$ lies in $(-1,1)$ since $p, s < 1$ and $p+s > 0$ at $\bar\beta$, so $\rho(M(\beta)) = 1$ exactly. The window is therefore open at its right endpoint and cannot be stated with $\le$; when $\bar\beta > 1$ the whole domain $(0,1)$ lies inside it and there is no admissible endpoint. □

Proof of Equation (11). In the two-timescale regime $\beta \ll 1-\rho_{\text{in}}$ we have $s-p = (1-\rho_{\text{in}}) - \beta(1-\rho_J+\rho_{\text{in}}) > 0$, of order $1-\rho_{\text{in}}$, while $qr = \beta^2\rho_{\text{in}}$. Expanding the exact root $\lambda_+ = \frac{p+s}{2} + \frac{s-p}{2}\sqrt{1+4qr/(s-p)^2}$ around $qr/(s-p)^2 \to 0$,

$$\begin{aligned}\lambda_+ &= s + \frac{qr}{s-p} + O\!\left(\frac{(qr)^2}{(s-p)^3}\right)\\ &= 1-\beta(1-\rho_J) + \frac{\beta^2\rho_{\text{in}}}{(1-\rho_{\text{in}})-\beta(1-\rho_J+\rho_{\text{in}})}\\ &\quad + O\!\left(\frac{\beta^4}{(1-\rho_{\text{in}})^3}\right),\end{aligned}$$

whose second term is $O\big(\beta^2/(1-\rho_{\text{in}})\big)$. Taking logarithms, the semi-log slope per iteration is $-\log\lambda_+ = \beta(1-\rho_J) + O(\beta^2/(1-\rho_{\text{in}}))$. □

The stated regime is not implied by the window bound: $s-p>0$ requires $\beta < \beta_c \coloneqq (1-\rho_{\text{in}})/(1-\rho_J+\rho_{\text{in}})$, and $\bar\beta \le \beta_c$ exactly when $\rho_{\text{in}} \ge (1-\rho_J)^2$; when $\rho_{\text{in}} < (1-\rho_J)^2$ the gap $\beta \in (\beta_c, \bar\beta)$ lies inside the window yet has $s-p<0$, where the slow eigenvalue is governed by $\rho_{\text{in}}(1+\beta)$.

Time-varying speeds inside the window (Definition 4.6) The recursion Equation (8) holds for every speed in $[0,1]$, so a single weighted norm turns Definition 4.6 into a guarantee for arbitrary schedules that stay below the window.

Corollary A.1 (Excited schedules)**.** *Assume $\eta(L^2-\tau^2) < 2\tau$ and let $0 \le \beta_t \le b$ for all t, with $0 < b < \bar\beta$ and $b \le 1$. Fix*

$$c \in \Big(\frac{1}{1-\rho_J},\; \frac{1-\rho_{\text{in}}(1+b)}{b\,\rho_{\text{in}}}\Big), \qquad \kappa \coloneqq \min\Big\{\frac{1-\rho_{\text{in}}}{b} - \rho_{\text{in}}(1+c),\; (1-\rho_J) - \frac{1}{c}\Big\} > 0, \tag{15}$$

(the interval is nonempty exactly when $b < \bar\beta$) and put $W_t \coloneqq e_t + c\,a_t$. Then for every t,

$$W_t \;\le\; W_0 \exp\Big(-\kappa \sum_{s<t}\beta_s\Big), \quad \text{NashConv}(z_t) \le L_{\text{gap}}\,\operatorname{dist}(z_t, Z^\star) \le L_{\text{gap}} \max\{1, \rho_J/c\}\, W_t. \tag{16}$$

Consequently (i) *if $\sum_t \beta_t = \infty$ then z_t converges to a point $z_\infty \in Z^\star$, with $\|z_t - z_\infty\| \le (1 + H/\kappa)\, W_t$ for $H = \max\{\rho_{\text{in}}, 1/c\}$;* (ii) *if (β_t) has windowed cumulative excitation with parameters (K, ε), then $W_t \le W_0\, e^{\kappa\varepsilon}\, \exp(-\kappa\varepsilon\, t/K)$, and the last iterate converges linearly to an exact Nash equilibrium of the original game.*

Proof. For $\beta \in [0, b]$, Equation (8) gives $W_{t+1} \le \big[\rho_{\text{in}}(1+\beta)+c\beta\rho_{\text{in}}\big]e_t + \big[\beta + c(1-\beta(1-\rho_J))\big]a_t$. The first bracket is at most $1-\kappa\beta$ if and only if $\beta\big(\rho_{\text{in}}(1+c)+\kappa\big) \le 1-\rho_{\text{in}}$, which holds on all of $[0,b]$ once $\kappa \le (1-\rho_{\text{in}})/b - \rho_{\text{in}}(1+c)$; this quantity is positive exactly when c is below the upper end of Equation (15). The second bracket is at most $c(1-\kappa\beta)$ trivially at $\beta = 0$ and, for $\beta > 0$, if and only if $1-c(1-\rho_J)+c\kappa \le 0$, i.e. $\kappa \le (1-\rho_J) - 1/c$, positive exactly when c exceeds the lower end. The two ends satisfy $1/(1-\rho_J) < (1-\rho_{\text{in}}(1+b))/(b\rho_{\text{in}})$ if and only if $b\rho_{\text{in}} < (1-\rho_J)(1-\rho_{\text{in}}-b\rho_{\text{in}})$,

i.e. $b\rho_{\text{in}}(2-\rho_J) < (1-\rho_{\text{in}})(1-\rho_J)$, which is $b < \bar{\beta}$ (and $b < \bar{\beta} < (1-\rho_{\text{in}})/\rho_{\text{in}}$ makes the upper end positive). Hence $W_{t+1} \le (1-\kappa\beta_t)W_t \le e^{-\kappa\beta_t}W_t$, and iterating gives Equation (16); the two distance bounds are the bridge of Proposition 4.4 with $e_t + \rho_J a_t \le \max\{1, \rho_J/c\}\, W_t$. For (i), the one-step inequality gives $\kappa\beta_t W_t \le W_t - W_{t+1}$, so $\sum_{s\ge t}\beta_s W_s \le W_t/\kappa$; the center drift obeys $\|\mu_{s+1}-\mu_s\| = \beta_s\|z_{s+1}-\mu_s\| \le \beta_s(\rho_{\text{in}}e_s + a_s) \le H\beta_s W_s$ (as in the proof of Proposition 4.4), so (μ_t) is Cauchy with $\|\mu_\infty - \mu_t\| \le HW_t/\kappa$, and $\mu_\infty \in Z^\star$ because $a_t \le W_t/c \to 0$ and $Z^\star$ is closed. Since $c > 1$, $\|z_t - \mu_t\| \le e_t + a_t \le W_t$ by Equation (14), hence $\|z_t - \mu_\infty\| \le (1+H/\kappa)W_t \to 0$. For (ii), $\sum_{s<t}\beta_s \ge \varepsilon\lfloor t/K\rfloor \ge \varepsilon(t/K - 1)$, and (i) with this geometric bound gives linear convergence to z_∞. □

At constant β the corollary recovers the conclusion of Theorem 4.5 with the cruder rate $e^{-\kappa\beta}$ in place of $\rho(M(\beta))$; its content is the schedule-free statement. The cap $b < \bar{\beta}$ cannot be dropped: on matching pennies at $(\eta, \tau) = (0.2, 0.5)$ the schedule $\beta_t \equiv 1$ has excitation with $K = 1$, yet the synchronized update $z_{t+1} = (I - \frac{\eta}{1+\eta\tau}F_0)z_t$ has skew-mode modulus $\sqrt{1 + (\eta\lambda/(1+\eta\tau))^2} > 1$, so the equilibrium is locally unstable (Appendix C).

Exactness of the resolvent factor (Lemma 4.3, part (b)) Part (b) is exact where it applies: matching pennies gives $\rho_J = 0.242536$ and the measurement returns it; rock–paper–scissors gives 0.277350. Part (a) covers boundary or non-unique equilibria: sampled values 0.665 on the random 3×4 game (unique boundary equilibrium, Appendix D) and 0.9325 on Kuhn (non-singleton equilibrium set; vertex-enumeration projection, independent of the dual-certificate routine of Appendix G).

B Proofs for Section 5

Proof of Theorem 5.1 (freezing)

Proof. *(i)* Each center step satisfies $\|\mu_{t+1} - \mu_t\| = \beta_t\|z_{t+1} - \mu_t\| \le \beta_t D$ since both points lie in Z. Summability of (β_t) makes (μ_t) a Cauchy sequence, so $\mu_t \to \mu_\infty$ with $\|\mu_\infty - \mu_0\| \le \sum_t \beta_t D = SD$.

(ii) Let $e_t = \|z_t - J(\mu_t)\|$ and $B_t \coloneqq D\beta_t$. By Lemma 4.1 (contraction of the inner step at frozen center) and nonexpansiveness of $\mu \mapsto J(\mu)$ (Lemma 4.2),

$$e_{t+1} \le \|z_{t+1} - J(\mu_t)\| + \|J(\mu_t) - J(\mu_{t+1})\| \le \rho_{\text{in}}e_t + \|\mu_{t+1} - \mu_t\| \le \rho_{\text{in}}e_t + B_t,$$

hence $e_t \le \rho_{\text{in}}^t e_0 + \sum_{k<t}\rho_{\text{in}}^{t-1-k}B_k$. The convolution of the geometric kernel with the summable sequence (B_k) vanishes: splitting the sum at $t/2$, the early part is bounded by $\rho_{\text{in}}^{\lfloor t/2\rfloor}\sum_k B_k$ and the late part by $(1-\rho_{\text{in}})^{-1}\sup_{k\ge t/2} B_k \le (1-\rho_{\text{in}})^{-1}\sum_{k\ge t/2} B_k$, both of which tend to zero. Thus $e_t \to 0$; combined with $\mu_t \to \mu_\infty$ and continuity of J, we get $z_t \to J(\mu_\infty)$, and continuity of $\mathrm{NashConv}$ gives the limit value. The convolution decays geometrically exactly when B_t itself does: if $B_t = O(q^t)$ then $e_t = O(\max\{\rho_{\text{in}}, q\}^t)$, with an extra factor t when $q = \rho_{\text{in}}$, which is why no rate is claimed in general. □

Proposition B.1 (Interior floor)**.** *Suppose the hypotheses of Theorem 5.1 hold with $S = \sum_t \beta_t$, the equilibrium z^{NE} is unique and relatively interior, the ball $B(\mu_0, SD)$ lies in the local linearization neighborhood of Section 6, and* $\mathrm{dist}(\mu_0, Z^\star) > SD$*. Then*

$$\liminf_t \mathrm{NashConv}(z_t) \;\ge\; c_0 \cdot \frac{\tau}{\sqrt{\tau^2 + \lambda_{\max}^2}} \cdot \big(\mathrm{dist}(\mu_0, Z^\star) - SD\big) \;>\; 0, \tag{17}$$

where $\lambda_{\max}$ is the largest skew strength of F on the tangent space and c_0 the local sharpness constant of $\mathrm{NashConv}$.

Proof of Proposition B.1 (interior floor)

Proof. In the local linearization neighborhood, tangent-space coordinates give (proof of Lemma 4.3(b))

$$J(\mu) - z^{\text{NE}} = \tau(\tau I + F_0)^{-1}(\mu - z^{\text{NE}}),$$

a normal linear map whose singular values are $\tau/|\tau + i\lambda_k| = \tau/\sqrt{\tau^2 + \lambda_k^2}$; its *smallest* singular value is $\tau/\sqrt{\tau^2 + \lambda_{\max}^2}$. By Theorem 5.1(i), $\|\mu_\infty - \mu_0\| \le SD$, so $\|\mu_\infty - z^{\mathrm{NE}}\| \ge \mathrm{dist}(\mu_0, Z^\star) - SD > 0$, and the whole segment from μ_0 to μ_∞ stays in $B(\mu_0, SD)$, inside the linearization neighborhood. Hence

$$\|J(\mu_\infty) - z^{\mathrm{NE}}\| \;\ge\; \frac{\tau}{\sqrt{\tau^2 + \lambda_{\max}^2}}\big(\mathrm{dist}(\mu_0, Z^\star) - SD\big).$$

Local sharpness of $\mathrm{NashConv}$ at the unique interior equilibrium, $\mathrm{NashConv}(z) \ge c_0\|z - z^{\mathrm{NE}}\|$ on the neighborhood with $c_0 > 0$ since $\mathrm{NashConv}$ is a polyhedral gap function vanishing only at z^{NE}, applied at $z = J(\mu_\infty) = \lim_t z_t$ gives Equation (17). □

The matching-pennies resolvent in closed form (Example 5.2) In the centered coordinates $u = 2p - 1 \in [-1, 1]$ and $v = 2q - 1$, the payoff operator of $A = \left[\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right]$ restricted to the simplex tangent line is $F_0(u, v) = (-2v, 2u)$, skew with strength $\lambda = 2$, and the unique equilibrium is $(0, 0)$, interior. The resolvent equation $0 = F_0(z^\star) + \tau(z^\star - \mu)$ is the linear system

$$\tau u^\star - 2v^\star = \tau\mu_u, \qquad 2u^\star + \tau v^\star = \tau\mu_v,$$

whose solution is $u^\star = (\tau^2\mu_u + 2\tau\mu_v)/(\tau^2 + 4)$, $v^\star = (\tau^2\mu_v - 2\tau\mu_u)/(\tau^2 + 4)$ as long as $(u^\star, v^\star)$ stays in the box (true near the equilibrium). The map has singular value $\tau/\sqrt{\tau^2 + 4}$ in every direction (it is a scaled rotation), so $\|(u^\star, v^\star)\| = \tau/\sqrt{\tau^2 + 4}\,\|\mu\|$. In these coordinates $Ay = (v, -v)$ and $A^\top x = (u, -u)$, so $\mathrm{NashConv}(u, v) = |v| - (-|u|) = |u| + |v| \ge \|(u, v)\|$, which gives the displayed lower bound. The frozen floors quoted in Example 5.2 are each computed by evaluating $\mathrm{NashConv}(J(\mu_\infty))$ at the trajectory's own μ_∞ and comparing with $\lim_t \mathrm{NashConv}(z_t)$; the identity holds trajectory by trajectory.

C Proofs for Section 6

Proof of Theorem 6.1 (interior local rate)

Proof. At a unique relatively interior equilibrium, no inequality constraint is active in a neighborhood, so Π_Z acts locally as the affine projection onto the affine hull, linear in coordinates centered at the equilibrium, and in those tangent-space coordinates the iteration Equation (4) is linear, with $F_0 = P_V F P_V$ the tangential compression of F onto the tangent space V (Lemma 4.3, part (b)):

$$z_{t+1} = \frac{z_t + \eta\tau\mu_t - \eta F_0 z_t}{1 + \eta\tau}, \qquad \mu_{t+1} = (1 - \beta)\mu_t + \beta z_{t+1}.$$

F_0 is skew and, by uniqueness, has no kernel on V (the perturbation argument of Appendix A); it decomposes into orthogonal invariant planes with eigenvalues $\pm i\lambda_k$, $\lambda_k > 0$. In the complex coordinate of mode k the pair (z, μ) evolves by the 2×2 system

$$z^+ = g\,z + h\,\mu, \qquad \mu^+ = \beta g\,z + \big(1 - \beta + \beta h\big)\mu, \qquad g = \frac{1 - i\eta\lambda_k}{1 + \eta\tau}, \quad h = \frac{\eta\tau}{1 + \eta\tau}.$$

At $\beta = 0$ the eigenvalues are g (fast, $|g| < 1$ under the step-size condition) and 1 (the center is invariant). Perturbing the unit eigenvalue in β: the quasi-static gain of the fast variable is

$$m_k \;=\; \frac{h}{1 - g} \;=\; \frac{\eta\tau/(1 + \eta\tau)}{\eta(\tau + i\lambda_k)/(1 + \eta\tau)} \;=\; \frac{\tau}{\tau + i\lambda_k},$$

and the slow eigenvalue is $1 - \beta(1 - m_k) + O(\beta^2)$, with modulus

$$\big|1 - \beta(1 - m_k)\big| = 1 - \beta\,(1 - \mathrm{Re}\, m_k) + O(\beta^2), \qquad \mathrm{Re}\, m_k = \frac{\tau^2}{\tau^2 + \lambda_k^2}.$$

The remainder is uniform over the finitely many modes, whose skew strengths lie in $[\lambda_{\min}, L]$. Since every fast eigenvalue has modulus $|g_k| + O(\beta)$ with $|g_k| < 1$ and every slow eigenvalue has modulus $1 - \beta(1 - \mathrm{Re}\, m_k) + O(\beta^2)$ with $\mathrm{Re}\, m_k$ largest at $\lambda_{\min}$, there is $\beta_{\mathrm{s}} > 0$, depending on the game and (η, τ), below which the slow eigenvalue of the $\lambda_{\min}$ mode dominates every eigenvalue

of every mode (a dominance threshold, generally distinct from a crossing of the two moduli). For $\beta < \beta_{\mathrm{s}}$ the worst-case asymptotic rate over initializations is the $\lambda_{\min}$ slow modulus; for constant β the mode systems are autonomous, so a given trajectory's rate is the dominant modulus among the eigenvectors on which its state expands with a nonzero coefficient, equal to the slow-mode modulus when the slow eigenvector of the $\lambda_{\min}$ mode carries a nonzero coefficient. This is Equation (12) with $\rho_J^2 = \tau^2/(\tau^2 + \lambda_{\min}^2)$; equivalently the decay slope $-\log r$ is $\beta\lambda_{\min}^2/(\tau^2 + \lambda_{\min}^2)$ to first order. The exponent transfers from the distance to the gap: $\mathrm{NashConv}$ is a polyhedral gap function vanishing only at z^{NE}, so $c_0\|z - z^{\mathrm{NE}}\| \le \mathrm{NashConv}(z) \le L_{\mathrm{gap}}\|z - z^{\mathrm{NE}}\|$ near the equilibrium (the sharpness constant $c_0 > 0$ of Proposition B.1), and $\log \mathrm{NashConv}(z_t)$ and $\log\|z_t - z^{\mathrm{NE}}\|$ differ by $O(1)$ at every nonzero iterate in that neighborhood, which leaves the asymptotic exponent unchanged; consecutive differences of $\log \mathrm{NashConv}$ need not converge, since the gap is piecewise linear and the mode rotates. □

The heuristic Equation (13) caps the center-limited slope $\beta(1 - \rho_J^2)$ at the inner budget $-\log \rho_{\mathrm{in}} \approx 1 - \rho_{\mathrm{in}}$ and intersects the two lines; no such cap is proved, and the formula is parameter dependent (at $\lambda = L = 0.1$ it returns $\beta^\star \approx 2.4$, outside $(0, 1]$, although the endpoint dynamics there are unstable). The exact picture is the characteristic polynomial of the mode system,

$$\xi^2 - \big(g + 1 - \beta + \beta h\big)\xi + g(1 - \beta) = 0,$$

whose two roots, followed continuously from $\xi_{\mathrm{s}}(0) = 1$ (slow) and $\xi_{\mathrm{f}}(0) = g$ (fast), have moduli $1 - \beta(1 - \mathrm{Re}\, m) + O(\beta^2)$ and $|g| - \beta|g|\,\mathrm{Re}\, m + O(\beta^2)$ (from $\xi_{\mathrm{f}}'(0) = -mg$): both decrease at first, and which branch eventually dominates depends on (λ, η, τ). On matching pennies at $(\eta, \tau) = (0.2, 0.5)$, $\lambda_{\min} = 2$, so $g = (10 - 4i)/11$, $h = 1/11$, $|g| = 0.979$; the fast modulus dips to 0.978 near $\beta \approx 0.05$ and then rises to $|g+h| = 1.064$ at $\beta = 1$, while the slow modulus falls monotonically to 0; the two cross at $\beta \approx 0.023$ (β_{s} for this game), and the moduli $(|\xi_{\mathrm{s}}|, |\xi_{\mathrm{f}}|)$ are $(0.971, 0.978)$ at $\beta = 0.03$, $(0.784, 0.9985)$ at $\beta = 0.2$, and $(0.673, 1.019)$ at $\beta = 0.3$. On this game the degrading rate past the crossover is the fast branch becoming dominant. The pattern depends on the game: at weak skew, $\lambda = 0.1$ at the same (η, τ), the roles reverse, the branch continued from g falling to 0 at $\beta = 1$ while the branch continued from 1 rises to modulus $|1 - i/55| > 1$. In either case the slow-mode modulus of Theorem 6.1 describes a trajectory only below β_{s}. The random 3×4 game of Section 6 (a unique boundary equilibrium) shows the same non-monotone shape: slope 2.06×10^{-2} at $\beta = 0.05$ but 4.9×10^{-3} at $\beta = 0.2$. As stated in Section 6, these are predictions of the extrapolated linear model.

Proof of Theorem 6.2 (tail rate of floored schedules) The precise form of the first claim: for every sufficiently small $\varepsilon > 0$, $[e; a]_t = O\big(\rho(\widetilde{M}_\varepsilon)^{\,t - T_0}\big)$, where $\widetilde{M}_\varepsilon$ elementwise dominates $M(\beta)$ over $\beta \in [\beta_{\min} - \varepsilon, \beta_{\min} + \varepsilon]$ and $\rho(\widetilde{M}_\varepsilon) \to \rho(M(\beta_{\min}))$ as $\varepsilon \downarrow 0$.

Proof. Fix $\varepsilon \in (0, \beta_{\min})$. Since $\beta_t \to \beta_{\min}$, there is T_0 with $\beta_t \in [\beta_{\min} - \varepsilon, \beta_{\min} + \varepsilon]$ for all $t \ge T_0$; the argument uses only this convergence, not monotonicity of the schedule. Define $\widetilde{M}_\varepsilon$ by taking the entries p, q, r of $M(\beta)$ at $\beta = \beta_{\min} + \varepsilon$ and the entry s at $\beta = \beta_{\min} - \varepsilon$; since p, q, r are nondecreasing and s is nonincreasing in β, this dominates $M(\beta_t)$ elementwise for every $t \ge T_0$. Iterating Equation (8), $[e; a]_t \le \widetilde{M}_\varepsilon^{\,t - T_0}\,[e; a]_{T_0}$ for $t \ge T_0$. The safety margin $\beta_t \le \bar{\beta} - \delta$ ensures, by continuity of the spectral radius in ε and Theorem 4.5, that $\rho(\widetilde{M}_\varepsilon) < 1$ for all sufficiently small ε, and $\rho(\widetilde{M}_\varepsilon) \to \rho(M(\beta_{\min}))$ as $\varepsilon \downarrow 0$, giving the claimed asymptotic exponent. Under the hypotheses of Theorem 6.1 the comparison bound drives the state into the linearization neighborhood, after which the local iteration is a linear difference system $x_{t+1} = A(\beta_t)x_t$ whose coefficients converge to $A(\beta_{\min})$, the $\beta_{\min}$ system of Theorem 6.1, not to the worst-case comparison matrix $\widetilde{M}_\varepsilon$. The tangent space splits orthogonally over the skew modes (the skew operator is normal; uniqueness removes any zero-strength mode), so the system decouples into complex 2×2 mode blocks $A_k(\beta_t)$. Let $\lambda_k^{\mathrm{s}}(\beta)$ and $\lambda_k^{\mathrm{f}}(\beta)$ be the eigenvalue branches continuing from 1 and from g_k at $\beta = 0$; $\lambda_k^{\mathrm{s}}(\beta) = 1 - \beta(1 - m_k) + O(\beta^2)$, so $|\lambda_k^{\mathrm{s}}(\beta)| = 1 - \beta(1 - \mathrm{Re}\, m_k) + O(\beta^2)$. Since $|g_k| < 1$ and, for $\lambda_k > \lambda_{\min}$, $1 - \mathrm{Re}\, m_k > 1 - \mathrm{Re}\, m_{\min}$ strictly while modes sharing $\lambda_{\min}$ have identical blocks, continuity over the finitely many modes gives a $\beta_{\mathrm{s}} \in (0, 1)$ such that for $\beta_{\min} < \beta_{\mathrm{s}}$ every block has $|\lambda_k^{\mathrm{f}}| < |\lambda_k^{\mathrm{s}}|$ and the $\lambda_{\min}$ blocks have the largest slow modulus; this is the smallness hypothesis of the theorem. The Perron-type theorem of Pituk (2002) for $x_{t+1} = (A + B_t)x_t$ with $B_t \to 0$ states

that every solution is eventually zero or has $\limsup_t \|x_t\|^{1/t}$ equal to the modulus of an eigenvalue of A; applied blockwise from a time T_0 after which $\beta_t < 1$, every mode solution not eventually zero has exponent $|\lambda_k^{\mathrm{s}}|$ or $|\lambda_k^{\mathrm{f}}|$. The solutions of a block with exponent at most $|\lambda_k^{\mathrm{f}}|$ form a complex linear subspace S_k (parameterized by the state at T_0), and $\dim_{\mathbb{C}} S_k \le 1$: if two independent solutions had that exponent, Hadamard's inequality would give $\limsup_t |\det \Phi_t|^{1/t} \le |\lambda_k^{\mathrm{f}}|^2$ for the fundamental matrix Φ_t, whereas $\det \Phi_t = \det \Phi_{T_0} \prod_{s=T_0}^{t-1} \det A_k(\beta_s)$ with $\det A_k(\beta_s) = g_k(1-\beta_s) \to g_k(1-\beta_{\min}) = \lambda_k^{\mathrm{s}}\lambda_k^{\mathrm{f}} \neq 0$ (here $\beta_{\min} < 1$ is used) gives $|\det \Phi_t|^{1/t} \to |\lambda_k^{\mathrm{s}}||\lambda_k^{\mathrm{f}}| > |\lambda_k^{\mathrm{f}}|^2$. Hence a mode trajectory outside S_k has exponent $|\lambda_k^{\mathrm{s}}|$, the exponent of the full state is the largest exponent over the modes, and when at least one $\lambda_{\min}$ block lies outside its S_k it equals the dominant modulus $1 - \beta_{\min}(1-\rho_J^2) + O(\beta_{\min}^2)$ with ρ_J the local factor of Lemma 4.3(b). The comparison matrix uses the global factor ρ_J^{gl} of Lemma 4.3(a); any global modulus satisfies $\gamma \ge 1/\lambda_{\min}$ (apply the subregularity inequality in the real $\lambda_{\min}$-invariant plane near the equilibrium, where $\mathrm{dist}(0, T(z)) = \|F_0\delta\| = \lambda_{\min}\|\delta\|$), so $\rho_J^{\mathrm{gl}} \ge \rho_J$ and $\rho(M(\beta_{\min})) = 1 - \beta_{\min}(1-\rho_J^{\mathrm{gl}}) + O(\beta_{\min}^2)$ exceeds the local modulus by at least $\beta_{\min}\rho_J(1-\rho_J)$ to first order; both expansions are as $\beta_{\min} \to 0$ with the game and (η, τ) fixed. □

Proof of Theorem 6.3 (vanishing speeds) The proof runs mode by mode in the 2×2 system of Appendix C. Fix a mode with skew strength λ, write g, h, $m = h/(1-g) = \tau/(\tau + i\lambda)$ as there, and change the fast variable to the mode's tracking error $\varrho := z - m\mu$; since $z_\tau^\star(\mu) = m\mu$ in the linear model, $e_t^2 = \sum_k |\varrho_k|^2$ and $a_t^2 = \sum_k |\mu_k|^2$. Using $gm + h = m$, the recursion becomes exactly

$$\varrho^+ = g(1 - m\beta_t)\,\varrho + m(1-m)\beta_t\,\mu, \qquad \mu^+ = \beta_t g\,\varrho + \big(1 - (1-m)\beta_t\big)\mu. \tag{18}$$

Write $w = 1 - m$, so $\mathrm{Re}\, w = 1 - \mathrm{Re}\, m = \lambda^2/(\tau^2 + \lambda^2) > 0$.

Proof. Invariance. Set $W = |\varrho| + K|\mu|$ with $K > |m||w|/\mathrm{Re}\, w$. From Equation (18), $|1 - \beta w| \le 1 - \beta\,\mathrm{Re}\, w + \beta^2|w|^2/2$, so $W^+ \le \big[|g|(1 + |m|\beta) + K\beta|g|\big]|\varrho| + \big[\beta|m||w| + K(1 - \beta\,\mathrm{Re}\, w + \beta^2|w|^2/2)\big]|\mu|$; both brackets are at most their coefficients in W once $\beta \le \beta_\circ$ with $\beta_\circ$ small, because $|g| < 1$ and $K\,\mathrm{Re}\, w > |m||w|$. The constants K, ρ, C, $\beta_\circ$ below are chosen mode by mode and then the least favorable over the finitely many modes is taken, so a single $\beta_\circ(R_0)$ serves all modes. Hence each mode's W_k is nonincreasing for $t \ge T$, and $U = \{\sum_k W_k < r\}$, a neighborhood of $(z^{\mathrm{NE}}, z^{\mathrm{NE}})$ in the (z, μ) state space, with r small enough that no inequality constraint is active, keeps the linear model exact along the whole trajectory. The map Equation (18) has determinant $g(1 - \beta_t) \neq 0$, so an excited mode stays nonzero, and the hypothesis $|\varrho_T| \le R_0|\mu_T|$ gives $\mu_T \neq 0$ on it; the bound $|\mu^+| \ge (1 - \beta(|w| + |g|r))|\mu| > 0$ below then keeps $\mu_t \neq 0$ for all $t \ge T$, so $a_t > 0$ whenever some mode is excited.

(i) Enslavement. Let $r_t = |\varrho_t|/|\mu_t|$ on an excited mode. From Equation (18), $|\mu^+| \ge (1 - \beta(|w| + |g|r))|\mu|$ and $|\varrho^+| \le |g|(1 + |m|\beta)|\varrho| + \beta|m||w||\mu|$. As long as $r_t \le R_0$ and $\beta_t \le \beta_\circ$, with $\beta_\circ$ small enough that $|g|(1 + |m|\beta_\circ) \le \rho\,(1 - \beta_\circ(|w| + |g|R_0))$,

$$r_{t+1} \;\le\; \rho\, r_t + C\beta_t, \qquad \rho := \tfrac{1+|g|}{2} < 1, \quad C := \frac{|m||w|}{1 - \beta_\circ(|w| + |g|R_0)}, \tag{19}$$

and the bound $r_t \le R_0$ propagates because $\rho R_0 + C\beta_\circ \le R_0$ once $\beta_\circ \le (1-\rho)R_0/C$, a condition met for small $\beta_\circ$ since C decreases to $|m||w|$ as $\beta_\circ \downarrow 0$. Iterating Equation (19), $r_t \le \rho^{t-T} r_T + C\sum_{s<t} \rho^{t-1-s}\beta_s \to 0$ because $\beta_s \to 0$. Since $e_t^2 \le \max_k r_{k,t}^2\, a_t^2$, this gives $e_t/a_t \to 0$.

(ii) Rate. On an excited mode, Equation (18) gives $\log|\mu_{t+1}| - \log|\mu_t| = \log\big|1 - \beta_t w + \beta_t g\,\varrho_t/\mu_t\big| = -\beta_t\,\mathrm{Re}\, w + O(\beta_t^2) + O(\beta_t r_t)$, with constants depending only on $|g|$, $|w|$, R_0. Summing, $-\log|\mu_t| = \mathrm{Re}\, w \sum_{s<t}\beta_s + O(\sum_{s<t}\beta_s^2) + O(\sum_{s<t}\beta_s r_s) + O(1)$, and both error sums are $o(\sum_{s<t}\beta_s)$ because $\beta_s \to 0$ and $r_s \to 0$. Since $z = \varrho + m\mu$ and $r_t \to 0$, $|z_{k,t}| = |m_k||\mu_{k,t}|(1 + o(1))$ on each excited mode, so $-\log|z_{k,t}|/\sum_{s<t}\beta_s \to 1 - \mathrm{Re}\, m_k$. When $\sum_t \beta_t = \infty$ every excited mode tends to zero, and the mode with the smallest exponent, $1 - \mathrm{Re}\, m$ at $\lambda_{\min}$, equal to $1 - \rho_J^2$ by Lemma 4.3(b), dominates every excited mode with $\lambda_k > \lambda_{\min}$ by a factor $\exp(-\Delta_k \sum_{s<t}\beta_s + o(\sum_{s<t}\beta_s)) \to 0$, where $\Delta_k = \mathrm{Re}\, m_{\min} - \mathrm{Re}\, m_k > 0$, and modes sharing $\lambda_{\min}$ share its exponent; since $\mathrm{dist}(z_t, Z^\star)$ is the norm of z_t in tangent coordinates, it is squeezed between the $\lambda_{\min}$ component and the sum over modes, hence $-\log \mathrm{dist}(z_t, Z^\star)/\sum_{s<t}\beta_s \to 1 - \rho_J^2$. A geometric rate $\mathrm{dist}(z_t, Z^\star) = O(\theta^t)$ with $\theta < 1$ would force $\sum_{s<t}\beta_s \ge c\,t$ for some $c > 0$, contradicting $\beta_t \to 0$.

(iii) Exact order. If $\sum_t \beta_t^2 < \infty$, then by Equation (19) and $\beta_t\beta_s \le (\beta_t^2+\beta_s^2)/2$, $\sum_t \beta_t r_t \le R_0 \sum_t \rho^{t-T} + C\sum_t\sum_{s<t}\rho^{t-1-s}\beta_t\beta_s \le \frac{R_0}{1-\rho} + \frac{C}{1-\rho}\sum_t \beta_t^2 < \infty$, so the error increments in (ii) are absolutely summable and, on each excited mode, $\log|\mu_{k,t}| + (1-\operatorname{Re} m_k)\sum_{s<t}\beta_s$ converges to a finite limit: $|\mu_{k,t}|\exp\big((1-\operatorname{Re} m_k)\sum_{s<t}\beta_s\big) \to C_k \in (0,\infty)$, with C_k depending on the trajectory through r_T, $\mu_{k,T}$, and the schedule; by $z_k = \varrho_k + m_k\mu_k$ and $r_t \to 0$ the same holds for $|z_{k,t}|$ with limit $|m_k|C_k$. Multiplying by $\exp((1-\rho_J^2)\sum_{s<t}\beta_s)$, an excited mode with $\lambda_k > \lambda_{\min}$ contributes $\exp(-\Delta_k\sum_{s<t}\beta_s)\,(|m_k|C_k + o(1)) \to 0$, while each excited $\lambda_{\min}$ mode contributes $|m_{\min}|C_k$; the mode planes are orthogonal, so

$$\operatorname{dist}(z_t, Z^\star)\,\exp\Big((1-\rho_J^2)\sum_{s<t}\beta_s\Big) \;\longrightarrow\; C_\star := \Big(\sum_{k:\ \lambda_k=\lambda_{\min}\ \text{excited}} |m_{\min}|^2 C_k^2\Big)^{1/2} \in (0,\infty),$$

which is the Θ statement with an asymptotic constant. For $\beta_t = c/(1+t)$ with $0 < c < 1$, $\sum_{s<t}\beta_s = c\log t + c\gamma_{\mathrm{E}} + o(1)$, so $\operatorname{dist}(z_t, Z^\star)\, t^{\,c(1-\rho_J^2)}$ converges to a positive constant, in particular $\operatorname{dist}(z_t, Z^\star) = \Theta(t^{-c(1-\rho_J^2)})$. If some mode is excited but the $\lambda_{\min}$ mode is not, (ii) and (iii) hold verbatim with $\lambda_{\min}$ replaced by the smallest excited skew strength. □

A quasi-static approximation, not a consequence of the bound Equation (19), complements the proof: freezing β_t and $q = \varrho_t/\mu_t$ over one step of the exact mode equation Equation (18) balances $q(1-g) = m(1-m)\beta_t$, i.e. to leading order in β, $|\varrho_t|/|\mu_t| \approx \beta_t\,|m(1-m)/(1-g)|$, which is $0.628\,\beta_t$ on matching pennies at $(\eta,\tau) = (0.2, 0.5)$; on the harmonic arm of Figure 2 the reproduced trajectory gives 6.28×10^{-5} at $t = 10^3$ ($\beta_t = 10^{-4}$) and 1.26×10^{-6} at $t = 5\times10^4$ ($\beta_t = 2\times10^{-6}$). The exponents $c(1-\rho_J^2)$ are 0.0941 on matching pennies ($\lambda_{\min} = 2$) and 0.0923 on rock–paper–scissors ($\lambda_{\min} = \sqrt{3}$); $\operatorname{dist}(z_t, Z^\star)\, t^{c(1-\rho_J^2)}$ is constant to four digits over the last quarter of the horizon (a longer window than the final-fifth $\mathrm{NashConv}$ tail slopes of Appendix D), and the least-squares slope of $\log\operatorname{dist}$ over that quarter, 2.1478×10^{-6} and 2.1065×10^{-6} per step, matches the slope of $-c(1-\rho_J^2)\log t$ on the same grid, 2.1477×10^{-6} and 2.1064×10^{-6}. The $\mathrm{NashConv}$ slopes of Appendix G sit within 5% and 0.6% of these values; $\mathrm{NashConv}$ is a piecewise-linear gap whose slope depends on the direction from which the slowly rotating mode approaches the equilibrium. Kuhn has a non-singleton equilibrium set, beyond the uniqueness hypothesis.

D Experimental Protocol

Carrier, games, and evaluation All layer-A dynamics run Equation (4) verbatim: a single projected proximal step per center update, simultaneous in the two players, followed by the center update using the *new* iterate. The projection onto the joint treeplex is the strictly convex quadratic program of Section 3, solved through its unconstrained dual by L-BFGS with duality-gap safeguards: the KKT residual is checked after every projection at tolerance 10^{-6} (relative to the input magnitude) and the run fails closed on violation; the dual gradient tolerance is 10^{-10} by default. Warm starts persist across iterations in both the iterate and the dual variables of the projection.

Unless stated otherwise, layer-A runs use $(\eta,\tau) = (0.2, 0.5)$, and both z_0 and μ_0 are the uniform realization plan. Horizons are 200k iterations on Kuhn and Leduc, and 100k and 50k on Liar's Dice with 4- and 5-sided dice, evaluated every $\max(100, T/500)$ iterations; the constant-β scans use 30k (Kuhn) and 50k (Leduc) with evaluation every 100; the Leduc rate runs extend to 400k with evaluation every 5000. $\mathrm{NashConv}$ is the full sequence-form saddle gap Equation (2), not divided by two, always of the *last* iterate. It is computed by an exact best response to the behavioral form of z_t, by backward induction over information sets ordered by depth and weighted by opponent and chance reach, summed over both players; the evaluator shares no code path with the dynamics' operator F. The floor 10^{-16} is applied when the values are reported.

Games All extensive-form trees are built by our own tabular engine (no external game library); A is the chance-weighted sequence-form payoff matrix. *Kuhn*: three cards, ante 1, one bet of size 1, standard rules; 6 information sets and 12 nonempty sequences per player. *Leduc*: three ranks in two suits, ante 1, two betting rounds with bet sizes 2 and 4, allowing at most two raises each round, player 1 acting first in both rounds, one public card from the remaining four; 144 information sets and 336 sequences per player. *Liar's Dice* with d-sided dice: each player rolls one die; bids

are pairs (quantity, face) with quantity in $\{1, 2\}$, increasing strictly in lexicographic order, no wild face; the "liar" call pays ± 1 according to whether at least the bid quantity of the bid face shows across both dice. The 4-sided game has 512 information sets and 1020 sequences per player, the 5-sided game 2560 and 5115. The matrix games are matching pennies $A = \left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right)$, the cyclic 3×3 rock–paper–scissors matrix, and a random 3×4 matrix with i.i.d. standard normal entries (NumPy `default_rng(0)`) rescaled to spectral norm $L = 2$; after iterated strict dominance it is a 2×2 game with a fully mixed equilibrium, so its equilibrium is unique and lies on the boundary, $x^\star \approx (0.777, 0.223, 0)$, $y^\star \approx (0.427, 0.573, 0, 0)$.

Arms and schedules The constant-β scans of Section 8 use the grid

$$\beta \in \{0.001,\ 0.002,\ 0.003,\ 0.005,\ 0.008,\ 0.01,\ 0.02,\ 0.05,\ 0.1,\ 0.2\}.$$

The excitation study of Section 5 runs seven schedules: $\beta_t = 0.1/(1+t)^2$ and $\beta_t = 0.1 \cdot 0.99^t$ (summable), $\beta \equiv 0$ (fixed center), pulses $\beta_t = \mathbf{1}[(t+1) \bmod K = 0]$ with $K \in \{50, 500\}$, the annealed floor $\beta_t = \max(0.1 \cdot 0.9999^t, 0.01)$, and the harmonic falsification arm $\beta_t = 0.1/(1+t)$, on all three carriers for 50k iterations, all seven arms sharing one random non-equilibrium start per matrix game (the uniform point is itself the equilibrium of matching pennies and of rock–paper–scissors and must be avoided). In the main comparison (Table 2), layer A comprises the fixed center, the moving centers $\beta = 0.05$ at $(\eta, \tau) = (0.2, 0.5)$ and at the original MMD parameters $(0.1, 1.0)$, and pulse re-centering with $K \in \{20, 50, 100, 200\}$; layer B comprises CFR$^+$, RTCFR$^+$, and Adp-RTCFR$^+$ as described in Appendix E.

The annealing study of Section 8 uses staircase schedules onto a floor, $\beta_t = \max(\beta_0\, d^{\lfloor t/s \rfloor}, \beta_{\min})$, with the four variants $(d, s) \in \{(0.8, 20000), (0.5, 5000)\}$ and $\beta_{\min} \in \{0.005, 0.01\}$, $\beta_0 = 0.05$ on Kuhn and 0.1 on Leduc, against the constant-β_0 control, all for 200k iterations evaluated every 400. The long-horizon extension runs the slow-decay Leduc variant $(0.8, 20000, 0.005)$ to 400k; it reaches its floor at iteration 280k. Its control is the constant speed $\beta = 0.005$, the floor value run from the start, which by 400k has not entered any asymptotic regime ($\mathrm{NashConv} = 2.15 \times 10^{-3}$, still falling), hence "pre-asymptotic."

The inner-step intervention of Section 7 runs m warm-started projected steps Equation (4) against a center that is updated once per block, on Leduc at $(\eta, \tau) = (0.2, 0.5)$ from the uniform start, for 50k center updates. With ρ_{in}^m in place of ρ_{in} the plug-in window endpoints are $\bar\beta_m = 0.0044,\ 0.0091,\ 0.0191,\ 0.0427,\ 0.1079$ for $m = 1, 2, 4, 8, 16$. At $\beta = 0.04$ the two arms inside their plug-in window ($m = 8, 16$) oscillate between 1.5×10^{-3} and 7×10^{-3} over the final 25k updates; at $\beta = 0.1$ the final NashConv falls monotonically in m, to 2.7×10^{-4} at $m = 8$ and 2.05×10^{-4} at $m = 16$ (about $5\times$ and $7\times$ below $m = 1$), while the local semi-log slope of the $m = 16$ arm, the only one inside its plug-in window, keeps flattening, its late-run value about a tenth of its early-run value: no arm shows an identifiable geometric regime over the horizon (the flattening slope is the observation).

Diagnostics and pre-registered criteria Curves are diagnosed on the evaluation grid with values floored at 10^{-16} (the evaluator's floating-point floor; raw values at or below zero occur only once the gap is numerically indistinguishable from zero at the evaluator's resolution). *Entry* is the first iteration with $\mathrm{NashConv} < 10^{-6}$ in the main comparison, and $\mathrm{NashConv} < 10^{-2}$ in the β-scans (whose purpose is ordering). The *asymptotic slope* is the semi-log regression slope over the range $10^{-12} < \mathrm{NashConv} < 10^{-5}$ (at least ten points; layer-A values below $\sim 10^{-11}$ sit at the projection-tolerance floor and the LP comparator of item (iii) below resolves 5×10^{-7}, so this is a descriptive fit, not a certified rate); the *tail slope* is the same regression over the final fifth of the curve, and the *tail coefficient of variation* and the *rebound* (final over minimum NashConv) flag oscillatory tails. For the fixed-speed and floored results the theoretical rate is the root exponent $r = \limsup_t \mathrm{dist}(z_t, Z^\star)^{1/t}$, the dominant eigenvalue modulus of the local system, and the positive numbers reported as slopes are decay slopes $-\log r$; for Theorem 6.3 the exponent is normalized by $\sum_{s<t} \beta_s$ rather than by t. Finite-window regression slopes and consecutive log differences are measured proxies for $-\log r$ and need not converge to it step by step, since rotating modes and polyhedral gap factors perturb them. For the excitation study two criteria were fixed before the runs: geometric behavior means either $\mathrm{NashConv}_T < 10^{-6}$ or a sustained tail slope above 10^{-5} per step, and the excitation indicator is the minimum window sum $\min_t \sum_{s=t}^{t+K-1} \beta_s$ at $K = 500$ with threshold 10^{-1}. Over the tested horizon the indicator is ≥ 1 on every geometric arm and at

most about 10^{-3} on the frozen and harmonic arms (1.0×10^{-3} for the harmonic arm at $T = 50$k, shrinking as t^{-1}), so the threshold separates the observed dichotomy. Each frozen arm is checked for tracking, $|\mathrm{NashConv}(z_T) - \mathrm{NashConv}(J(\mu_T))| \le 0.1 \max(\mathrm{NashConv}(J(\mu_T)), 10^{-9})$.

For Section 7, the error-bound modulus is measured directly from the geometry. On the sequence-form carrier we sample radial points between the equilibrium set and random Dirichlet directions, 73 on Leduc and 25 on Kuhn, and at each point compute the true manifold distance $\mathrm{dist}(z, Z^\star)$ by quadratic-programming projection onto $X^\star \times Y^\star$ together with the tangent-cone residual $\mathrm{dist}(0, F(z) + N_Z(z)) = \|\Pi_{T_Z(z)}(-F(z))\|$, with $\Pi_{T_Z(z)}$ the projection onto the tangent cone of Z at z; the modulus estimate is the worst-case ratio between the two. The result is $\gamma \approx 4.56$ on Leduc (99th percentile 4.13, median 0.62; 15 of the 73 ratios exceed 3) against $\gamma \approx 4.16$ on Kuhn (99th percentile 4.15, median 0.95). Both projections write each side of $Z^\star$ through its dual variables, as in Appendix G; the distances are interior-point solutions, not certified nearest points, so the modulus estimate is numerical evidence.

Proxy definitions and initial states Two proxy families appear, and only the manifold-distance samples above bound the modulus of Lemma 4.3. On Leduc the single-point proxy $\rho_J = 0.9117$ is the sample maximum, over 64 random points μ (Gaussian perturbations of scale 3 projected onto the treeplex, seed 0), of $\|J(\mu) - z^{\mathrm{ref}}\|/\|\mu - z^{\mathrm{ref}}\|$, where z^{ref} is one LP equilibrium and $J(\mu)$ the resolvent warm-started at μ. It is a heuristic diagnostic: a point of $Z^\star$ other than z^{ref} has positive distance to z^{ref} and zero residual, so this ratio does not lower-bound the modulus. On Kuhn the value $\rho_J = 0.9325$ is instead the sample maximum over 80 points of $\mathrm{dist}(J(\mu), Z^\star)/\mathrm{dist}(\mu, Z^\star)$ with the quadratic-programming projection onto $X^\star \times Y^\star$; it samples the manifold-distance contraction of J, and differs from $\rho_J^{\mathrm{eff}} = 0.901$ in the quantity sampled (a contraction ratio of J against the tangent-cone residual ratio γ). Throughout, $\rho_{\mathrm{in}} = \sqrt{1 + \eta^2 L^2}/(1 + \eta\tau)$. For the matrix-game rate comparisons of Section 6 at $\beta = 0.01$: $(\eta, \tau) = (0.2, 0.5)$, $T = 60$k, $\mu_0 = z_0$, with $z_0 = (0.5467, 0.4533;\ 0.3998, 0.6002)$ on matching pennies and $z_0 = (0.5446, 0.4478, 0.0076;\ 0.2315, 0.3672, 0.4013)$ on rock–paper–scissors (the uniform point is the equilibrium). The measured slope is the unweighted semi-log regression of $\|z_t - z^\star\|$ over $10^{-12} < \|z_t - z^\star\| < 10^{-5}$ (iterations 867–2561 and 1005–2730; 1695 and 1726 points), returning 9.5132×10^{-3} and 9.3414×10^{-3}. Over the same protocol windows the $\mathrm{NashConv}$ slopes are 9.4857×10^{-3} (iterations 907–2632; 1726 points) and 8.9900×10^{-3} (1050–2801; 1752 points). The random 3×4 game (generator of Appendix D) starts at $z_0 = (0.2005, 0.7995, 0;\ 0.1513, 0.0896, 0.1657, 0.5934)$ and returns 5.653×10^{-3} over iterations 1949–4799 (2851 points); its $\mathrm{NashConv}$ slope over the same window is 5.616×10^{-3}. Its $\rho_J = 0.665$ is the sample maximum of $\|J(\mu) - z^\star\|/\|\mu - z^\star\|$ over 120 points ($z^\star$ unique).

Determinism and numerical controls Every dynamic in the paper is tabular and sample-free. Determinism was verified directly: three seeds of the CFR$^+$ baseline on Kuhn produce bitwise-identical $\mathrm{NashConv}$ curves, so all reported numbers are single runs. Two controls delimit the small numbers. First, tightening the projection's dual gradient tolerance from 10^{-10} to 10^{-12} (with triple the iteration budget) lowers the layer-A Kuhn floors from $\sim 10^{-11}$ by 18–100$\times$; those floors come from projection tolerance, while every reported Leduc value is bitwise invariant to the same change: the finite-horizon behavior of Section 7 is insensitive to this tolerance. Second, the long-horizon extensions (500k Leduc, 400k floor runs, 200k on Liar's Dice with 5-sided dice) reproduce their parent runs digit for digit at all shared evaluation grid points, which is the regression test guarding every extension reported in Section 8.

Independent validation of the engine and evaluator Four checks test the game construction and the gap evaluator against independent computations. (i) The sequence-form linear program (Von Stengel, 1996; HiGHS) on our (E, e, A) returns the game values $-1/18$ on Kuhn (exact), -0.085606 on Leduc against OpenSpiel's -0.085603 ± 10^{-4} (its CFR$^+$ average after 2000 iterations, $\mathrm{NashConv} = 1.7 \times 10^{-4}$), and $-1/8$ on 4-sided Liar's Dice; with suit isomorphism enabled, OpenSpiel's Kuhn and Leduc have 12 and 288 information sets, equal to ours. (ii) On nine random non-equilibrium policies per game (Gaussian perturbations of the uniform point projected onto the treeplex, $\mathrm{NashConv}$ between 0.6 and 13), the tree-recursion evaluator and an LP best response agree to a relative 1.6×10^{-8} (the LP tolerance) on Kuhn, Leduc, and 4-sided Liar's Dice. (iii) Re-running RTCFR$^+$ on Leduc at its published $(\mu_{\mathrm{rt}}, T_u) = (0.001, 100)$ and evaluating the perturbed last iterate every 500 iterations, unclipped, by both evaluators: the values agree to the LP's absolute tolerance

($\leq 5 \times 10^{-7}$; relative 10^{-3} or better above 10^{-4}), stay positive throughout, the iterate's value $x^\top A y$ equals the game value to six digits from iteration 1000 on, flow residuals are $\leq 2 \times 10^{-16}$, and the entry iteration ($\mathrm{NashConv} = 9.1 \times 10^{-7}$ at 6000) matches Table 2. On 4-sided Liar's Dice both evaluators return exactly 0 from iteration 250 on (1.1×10^{-15} and 2.5×10^{-16} at 200), corroborating the rapid layer-B entries of Table 2 in kind. (iv) Independent evaluation in OpenSpiel (Lanctot et al., 2019). Each of our games is matched to an OpenSpiel game by a synchronous depth-first traversal of the two trees that asserts, node by node, the acting player, the number of actions, the chance probabilities, and the terminal payoffs, and records the information-set correspondence, which the traversal asserts to be a bijection between the two partitions with identical legal-action lists on each pair; no information-set string is parsed. Kuhn maps to `kuhn_poker`; Leduc to `leduc_poker` with suit isomorphism (its default partition, 936 sets, refines ours and is not used); Liar's Dice to a clean-room OpenSpiel game that we register from the rules of Appendix D, because the built-in `liars_dice` treats the highest face as wild; the clean-room game is an independent implementation of those rules. On the uniform policy, 30 Dirichlet(1) policies, and CFR$^+$ iterates at 200 iterations (2000 on Kuhn and Leduc), our evaluator and twice OpenSpiel's exploitability agree to 8.9×10^{-16} on all four games. For each of the twelve layer-B entries of Table 2 the arm is then re-run, the evaluated policy $\hat{\sigma}$ is saved at the powers of two, near the entry iteration, at half the horizon, and at the horizon (7 to 23 checkpoints per entry, 224 in all), the re-run is checked to reproduce the logged final, minimum, and entry values digit for digit, and every saved policy is evaluated in OpenSpiel. The largest absolute difference over the 224 policies is 8.9×10^{-16} (Leduc CFR$^+$ at $\mathrm{NashConv} = 3.4 \times 10^{-3}$, relative 2.6×10^{-13}) and 4.4×10^{-16} over the RTCFR$^+$ and Adp-RTCFR$^+$ entries, 5-sided Liar's Dice included; at the horizon OpenSpiel returns 1.635×10^{-14} and 1.654×10^{-14} for RTCFR$^+$ and Adp-RTCFR$^+$ on Leduc and 0 to 10^{-16} on both Liar's Dice sizes. These checks cover the game construction, the gap evaluator, and the evaluated policies of every layer-B entry; the update rule of RTCFR$^+$ is checked in Appendix E.

E Reproduction Details for the Layer-B Baselines

CFR$^+$. Tabular CFR$^+$ with alternating updates (player 0's regrets and strategies are updated over all information sets, then player 1's), regret matching$^+$ ($R \leftarrow \max(R + \hat{r}, 0)$ with uniform fallback at zero mass), and linearly weighted strategy averaging. Table 2 reports the *last iterate's* $\mathrm{NashConv}$, matching the metric of every other arm; the linear average converges as expected.

RTCFR$^+$. Our runs use the published hyperparameters verbatim: regularization weight and reset period $(0.1, 10)$ on Kuhn, $(0.001, 100)$ on Leduc, $(0.01, 10)$ for Liar's Dice with 4-sided dice, and $(0.0005, 10)$ on Liar's Dice with 5-sided dice; write them (μ_{rt}, T_u). With $\gamma_0 = 10^{-10}$, the perturbed strategy at information set I with $|A_I|$ actions is $\hat{\sigma}_I = (1 - \gamma|A_I|)\sigma_I + \gamma$. At the start of every round t with $t \bmod T_u = 0$ the weight is halved, $\gamma \leftarrow \gamma/2$, and the reference is set to the sequence form r of the current $\hat{\sigma}$. Within a round, players alternate; for each I of the updating player, with counterfactual values q_I and own reach π_I,

$$g_I = q_I - \mu_{\mathrm{rt}}\big(\pi_I \hat{\sigma}_I - r_I\big), \qquad \theta_I \leftarrow \big[\theta_I + g_I - \langle \sigma_I, g_I \rangle\big]_+, \qquad \sigma_I \propto \theta_I$$

(uniform when $\theta_I = 0$); the cumulative regrets θ persist across resets, and $\mathrm{NashConv}$ is evaluated on $\hat{\sigma}$. The reproduction is checked before use: on Leduc the implementation reaches $\mathrm{NashConv} = 1.511 \times 10^{-14}$ within 20k iterations, matching the reference implementation's reported gap, and this check is asserted programmatically before the main runs. The update itself is compared with an independent re-implementation written from the released implementation of Meng et al. (2025) and run on the OpenSpiel trees of Appendix D. That implementation propagates the regularized counterfactual values upward through the tree, whereas Algorithm 1 of Ren et al. (2025) (line 16) subtracts the regularizer once, at the information set; our implementation follows the latter form. Against the re-implementation in that form the regularized counterfactual gradients computed from identical supplied states (five random pairs of $\hat{\sigma}$ and r) agree to 3.3×10^{-16} on Kuhn, Leduc, and 4-sided Liar's Dice, and the trajectories to 1.9×10^{-10} over 500 iterations on Kuhn and 2.9×10^{-7} over 300 on 4-sided Liar's Dice; on Leduc the trajectories separate at iteration 2 at one information set whose cumulative regret is 3.5×10^{-21}, a tie that regret matching$^+$ resolves differently under the two summation orders, while the paired-state gradients agree throughout. The propagating form departs from ours by $O(1)$ in $\hat{\sigma}$ from iteration 2 on; run under the evaluation grid of Table 2 it enters

$\mathrm{NashConv} < 10^{-6}$ at iterations 400, 5600, 200, and 1000 on Kuhn, Leduc, and 4- and 5-sided Liar's Dice, against 400, 6000, 200, and 300 for the form we run; both forms reach the floating-point floor.

Adp-RTCFR$^+$. A good-faith tabular reimplementation of the published Algorithm 1 (no tabular reference release): the reset period of RTCFR$^+$ is replaced by a three-state trigger on measured exploitability. The inner update is regret matching$^+$ on $g_I = q_I - w\mu_{\mathrm{rt}}(\sigma_I - \sigma_I^r)$, a behavioral-strategy deviation from the reference σ^r (initially uniform) with adaptive weight w (initially 1), without discounting, players alternating. After every round the full $\mathrm{NashConv}(\sigma)$ is computed and compared with its running minimum $\epsilon_{\min}$ over a counter k of player passes: if $\mathrm{NashConv} \le \epsilon_{\min}/2$, the reference is refreshed, $\epsilon_{\min}$ updated, $w \leftarrow 2$, $k \leftarrow 0$ (exploit); else if $\mathrm{NashConv} \le \epsilon_{\min}$ and $k \ge T_u$, the same refresh with $w \leftarrow 1$ (keep); else if $k \ge 2T_u$, the reference is refreshed without updating $\epsilon_{\min}$ and $w \leftarrow 0.5$ (explore). Regrets persist across refreshes; (μ_{rt}, T_u) are taken from the RTCFR$^+$ table above, not from the Adp-RTCFR$^+$ values of Ren et al. (2025, Table 3), which are $(0.05, 5)$ on Kuhn and $(0.01, 200)$ on Leduc and give no value for 4- or 5-sided dice; there the counter k and the period T count player passes, whereas T_u above counts full iterations. A step-by-step comparison with Algorithm 1 finds no other deviation; the alternation of players is our choice where the pseudocode leaves the order open.

F Reading Table 2

Four regularities organize Table 2, read as iteration counts and final gaps (trajectories in Figure 1). The moving center converges wherever any layer-A method does, earliest on Kuhn and within $1.3\times$ on Liar's Dice, and leads re-centering on Leduc at long horizons; entry grows monotonically in K. The fixed center freezes at game-scale values (Section 5). The two RTCFR variants of layer B reach the floating-point floor on all four games (Leduc 1.7×10^{-14}; CFR$^+$ stays at 10^{-3} on Kuhn and Leduc) while layer A sits at 10^{-4} on Leduc.

Scope of the guarantees The theorems hold for the Euclidean carrier-matched proxies; native MMD, APMD, and RTCFR$^+$ enter only through them, and no rate transfers to their native geometries. On the extensive-form games the window is a sampled plug-in, so a speed inside it carries no certificate; the $m = 16$ failure traces to the plug-in ρ_J (Section 7), and the layer-B gaps reproduce to 9×10^{-16} in OpenSpiel (Appendix D).

G Deferred Data: The τ Scan and the Excitation Study

Theorem 4.5 bounds the window by the two structural factors ρ_{in} and ρ_J, and Section 7 reports the τ scan that separates them on Leduc; the tables below are cited from the main text.

Annealing onto floors (deferred from Section 8) These runs compare schedules; neither game satisfies the interior hypothesis of Theorem 6.2. On Kuhn all four annealing variants enter at the same iteration as the constant-β_0 control (several coincide with it before entry by construction); on Leduc annealing is $3\times$ to $21\times$ worse at 200k. The slow-decay Leduc arm reaches 1.07×10^{-4} by 400k against 2.15×10^{-3} for the constant β=0.005 control. On Kuhn both land at the projection-tolerance floor ($\approx 2 \times 10^{-11}$).

Constant schedules outside the estimated window On Kuhn the six arms with β beyond the estimated window also converge, with slopes saturating at 5–8 $\times 10^{-4}$ per step, far below the inner budget $1 - \rho_{\mathrm{in}} \approx 0.08$. Leaving the window buys earlier entry, with the tail slope unchanged.

τ	ρ_{in}	$\rho_J(\gamma{=}4.56)$	$\bar{\beta}^{\mathrm{eff}}$
0.5	0.9485	0.916	0.00422
1.0	0.8695	0.977	0.00341
2.0	0.7453	0.994	0.00203

Table 3: Scanning τ on Leduc with the radial modulus $\gamma = 4.56$.

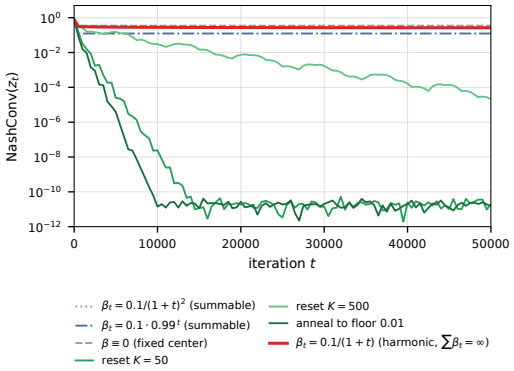


Figure 2: Seven schedules on Kuhn poker (50k).

β	$\rho(M(\beta))$	$1 - \rho(M(\beta))$
0.003 (inside est. window)	0.999911	8.9×10^{-5}
0.005	1.000065	-6.5×10^{-5}
0.010	1.001306	-1.3×10^{-3}
0.020	1.007671	-7.7×10^{-3}
0.050	1.044451	-4.4×10^{-2}
0.100	1.118085	-1.18×10^{-1}
0.200	1.270063	-2.70×10^{-1}

Table 4: Spectral radius of the tracking matrix $M(\beta)$ on Leduc at the operating $\rho_{\text{in}} = 0.9485$, $\rho_J = 0.9117$.

The tables below are calibrated on the radial moduli γ of Appendix D: on Leduc the factor $\rho_J^{\text{eff}} = 0.916$ sits next to the single-point proxy $\rho_J = 0.9117$, and on Kuhn the calibrated $\rho_J^{\text{eff}} = 0.901$ widens the window from $\bar{\beta} = 0.0057$ to 0.0082, a window the sampled ratio 0.9325 keeps closed. Along the $m = 16$ run the measured ratio exceeds every radial calibration from $t = 7$ on (Appendix G). At $\beta = 0.003$ the quadratic term $\beta^2 \rho_{\text{in}}/(1 - \rho_{\text{in}}) = 1.66 \times 10^{-4}$ and the linear gain 2.65×10^{-4} give a comparison-matrix rate of $\approx 8.9 \times 10^{-5}$ per step. On Leduc the true manifold distance falls to 3.8×10^{-3} by 200k while $\|z_t - z^{\text{ref}}\|$ stays at 1.54, where z^{ref} is the single equilibrium (an LP solution) at which the single-point proxy ρ_J is evaluated: the iterate approaches a point of $Z^\star$ other than z^{ref}, the same signature as Kuhn, where $\|z_t - z^{\text{ref}}\| \to 1.13$ while $\text{dist}(z_t, Z^\star) \to 9 \times 10^{-10}$.

The excitation dichotomy of Section 5 is carried by Figure 2: seven schedules on Kuhn poker (50k), grouped by excitation structure. The summable schedules ($\beta_t = 0.1/(1+t)^2$, $\beta_t = 0.1 \cdot 0.99^t$, and $\beta \equiv 0$) all freeze at $\text{NashConv}(J(\mu_\infty))$, at 0.3436, 0.1237, and 0.3457 on Kuhn; the geometric-decay schedule stops near the equilibrium on the matrix games ($\text{NashConv} \leq 3.5 \times 10^{-5}$). The windowed-excitation schedules (period-50 and period-500 resets; annealing onto a floor) converge geometrically on all carriers: final NashConv down to 1.94×10^{-11} on Kuhn and 10^{-16} on the matrix games at period 50, and a tail slope of 1.84×10^{-4} per step at period 500. The falsification arm ($\beta_t = 0.1/(1+t)$, harmonic) has $\sum_t \beta_t = \infty$ yet vanishing window sums; its dynamics end at $\text{NashConv} = 2.63 \times 10^{-1}$ on Kuhn ($\leq 1.36 \times 10^{-1}$ on the matrix games), nine to ten orders above the geometric arms, with tail slopes 0.9–2.1 $\times 10^{-6}$ per step, below the protocol's geometric criterion of 10^{-5}; on matching pennies and rock–paper–scissors the tails agree with the power laws predicted by Theorem 6.3 (slopes matched in Appendix C; the theorem's trajectory-entry hypotheses were not certified on these runs); the random matrix game has a boundary equilibrium, beyond that theorem's interior hypothesis.

Which input of the $m = 16$ plug-in fails The two rows of Proposition 4.4 are checked directly along the $\beta = 0.04$, $m = 16$ Leduc run of Section 7, re-run with the state saved at 24 checkpoints ($t \leq 10$, then 20, 50, 100, …, 49,999); its final NashConv lies between 1.4×10^{-3} and 6.6×10^{-3}

over the last 25k updates, as in the original run. At each checkpoint the resolvent $J(\mu_t)$ is the fixed point of the inner step at frozen center, iterated to a step change of 10^{-13}; one further inner step moves it by at most 1.6×10^{-7}, and the projection error of that step is at most 1.9×10^{-7}. The inner step contracts by $\rho_{\text{in}} = 0.9485$ at frozen center (Lemma 4.1), so the computed resolvent lies within $(1.6 + 1.9) \times 10^{-7}/(1 - \rho_{\text{in}}) \leq 6.8 \times 10^{-6}$ of the exact one at every checkpoint. This error is at most 17% of the smallest tracking error $e_t = \|z_t - J(\mu_t)\| = 4.2 \times 10^{-5}$. Since $\rho_J^{\text{emp},t} = \text{dist}(J(\mu_t), Z^\star)/a_t$, it moves $\rho_J^{\text{emp},t}$ by at most $6.8 \times 10^{-6}/a_t$, which is 7.0×10^{-5} at the final checkpoint's $a_t = 0.098$ and smaller wherever a_t is larger. Holding the center distances fixed, it contributes at most $\beta\rho_{\text{in}}^m \cdot 6.8 \times 10^{-6} \leq 1.2 \times 10^{-7}$ to a center-row slack with a fixed ρ_J; in the empirical-factor column the numerator $\text{dist}(J(\mu_t), Z^\star)$ carries the same error, giving at most $\beta(1 + \rho_{\text{in}}^m) \cdot 6.8 \times 10^{-6} \leq 3.9 \times 10^{-7}$ from the resolvent error alone. Neither estimate includes the uncertainty in the computed center distances. The slacks in Table 5 are reported as computed, without a certified error interval: the resolvent error above is a contraction bound on the computed fixed point, the distances below are interior-point solutions checked for equilibrium membership rather than nearest-point optimality, and the error propagation was not recorded per checkpoint. The diagnostic is numerical evidence for the failure of the plug-in ρ_J, not a rigorous certificate; every negative $s_c(0.9117)$ among the displayed checkpoints has magnitude at least 3.4×10^{-4}, three orders above the resolvent-error contribution. The distances $a_t = \text{dist}(\mu_t, Z^\star)$ and $\text{dist}(J(\mu_t), Z^\star)$ are Euclidean distances in sequence form to $X^\star \times Y^\star$, each factor written through its dual certificate ($x \in X^\star$ iff some u satisfies $E_1^\top u \leq A^\top x$ and $e_1^\top u \geq v^\star$; $y \in Y^\star$ iff some w satisfies $E_0^\top w \geq Ay$ and $e_0^\top w \leq v^\star$) and solved by interior-point quadratic programming (tolerance 10^{-12}); every projected point is re-certified by an independent linear-program best response to 1.1×10^{-10}, and on Kuhn the routine agrees with a vertex-enumeration projection over 20 points to a relative 9.2×10^{-9}. Since J is firmly nonexpansive with fixed-point set $Z^\star$, the exact ratio $\rho_J^{\text{emp},t} = \text{dist}(J(\mu_t), Z^\star)/a_t$ cannot exceed 1, and the computed values respect this necessary condition (Table 5). The tracking row $e_{t+1} \leq \rho_{\text{in}}^m(1 + \beta)e_t + \beta a_t$ holds at all 24 checkpoints with slack of at least 3.8×10^{-3}; the measured inner contraction κ_{in}, the ratio of the tracking error after the m inner steps to e_t, is at most 0.229 against the budget $\rho_{\text{in}}^{16} = 0.429$; projection and replay errors are at most 1.9×10^{-7}. The resolvent error of 6.8×10^{-6} enters both terms of κ_{in} and keeps it below the budget at every displayed checkpoint except the last, where $e_t = 4.2 \times 10^{-5}$ is within a factor of seven of that error and the worst case $(0.229\, e_t + 6.8 \times 10^{-6})/(e_t - 6.8 \times 10^{-6}) = 0.47$ exceeds it. The center row $a_{t+1} \leq (1 - \beta(1 - \rho_J))a_t + \beta\rho_{\text{in}}^m e_t$ holds at all 24 checkpoints with the nonexpansive value $\rho_J = 1$ and with the measured $\rho_J^{\text{emp},t}$ substituted; it fails at 12 of the 24 (all $t \geq 50$) with the plug-in $\rho_J = 0.9117$ and with the radially calibrated 0.916, and at 23 with 0.800. The measured ratio rises from 0.906 at $t = 0$ past 0.9117 at $t = 4$ to 0.987 at $t = 1000$ and stays between 0.9977 and 0.9994 from $t = 5000$ on; it is unchanged to four digits when the optimal-set tolerance is set to 10^{-10} or 0, the quadratic-program tolerance to 10^{-12} or 10^{-14}, and the resolvent step tolerance to 10^{-11} or 10^{-13}. Through Equation (7) at $\tau = 0.5$, 0.987 corresponds to $\gamma \geq 12$ and 0.9994 to $\gamma \geq 57$, against the radial sample $\gamma \approx 4.56$ of Appendix D: along this trajectory the modulus is far larger than the radial samples indicate, the same direction as the Kuhn contraction sample 0.9325. The computed diagnostics point to the resolvent contraction constant of the center row as the failing input, rather than the inner contraction or the update error. With late-run values between 0.9977 and 0.9994 in place of 0.9117 the plug-in matrix has $\rho(M_{16})$ between 1.0011 and 1.0012 instead of 0.9977, so the comparison bound certifies no decay; the dynamics do decay, slowly, with measured per-block ratios of 0.99994 for e_t and 0.99995 for a_t over $t = 5000$ to 40,000. The bound is loose along this run: at the last checkpoint its tracking row allows $e_{t+1} \leq 3.94 \times 10^{-3}$ where the measured value is 4.2×10^{-5}. The diagnosis indicates the failing input, not the rate, and as numerical evidence it does not certify the exclusion of the alternatives.

An earlier version of this diagnostic used a defective certificate for $Y^\star$: the dual inequalities of the minimizing player were written with the opposite sense, every $y \in Y$ satisfied them, $Y^\star$ degenerated to Y, and a_t measured $\text{dist}(\mu_{x,t}, X^\star)$ alone. Ratios computed with that distance exceeded 1 (up to 1.0101 at $t = 40{,}000$), which a firmly nonexpansive map cannot produce. The certificate above is the corrected one, checked on Kuhn against a vertex-enumeration projection over 20 points to a relative 9.2×10^{-9} (the defective one differs from it by 79%). Every quantity downstream of the optimal-set distance was recomputed: the sampled moduli (γ on Leduc 2.67 $\to$ 4.56, on Kuhn 3.11 $\to$ 4.16), the calibrated factors (0.800 $\to$ 0.916, 0.841 $\to$ 0.901), the estimated windows (0.009 $\to$ 0.0042, 0.012 $\to$ 0.0082), Table 3, and the Leduc manifold distance at 200k ($1.7 \times 10^{-3} \to 3.8 \times 10^{-3}$). The

Kuhn contraction sample 0.9325 and the random-game value 0.665 come from vertex enumeration and are unaffected. The 24 checkpoint records and the auditing script are in the code package.

t	$\mathrm{NashConv}(z_t)$	e_t	a_t	$\rho_J^{\mathrm{emp},t}$	κ_{in}	s_{tr}	$s_{\mathrm{c}}(0.9117)$	$s_{\mathrm{c}}(1)$	$s_{\mathrm{c}}(\mathrm{emp})$
0	4.75	1.73	5.61	0.9057	0.211	0.606	0.034	0.054	0.033
10	1.77	0.051	5.35	0.9211	0.221	0.186	5.2×10^{-3}	0.024	7.2×10^{-3}
100	0.619	0.021	4.32	0.9703	0.219	0.161	-8.9×10^{-3}	6.3×10^{-3}	1.2×10^{-3}
1,000	0.164	3.8×10^{-3}	2.56	0.9869	0.218	0.100	-7.6×10^{-3}	1.4×10^{-3}	8.3×10^{-5}
5,000	0.023	5.2×10^{-4}	0.811	0.9977	0.218	0.032	-2.8×10^{-3}	8.4×10^{-5}	1.0×10^{-5}
10,000	0.011	2.4×10^{-4}	0.458	0.9984	0.218	0.018	-1.6×10^{-3}	3.4×10^{-5}	4.8×10^{-6}
40,000	3.0×10^{-3}	6.4×10^{-5}	0.135	0.9992	0.218	5.4×10^{-3}	-4.7×10^{-4}	5.5×10^{-6}	1.3×10^{-6}
49,999	2.5×10^{-3}	4.2×10^{-5}	0.098	0.9994	0.218	3.9×10^{-3}	-3.4×10^{-4}	3.1×10^{-6}	8.3×10^{-7}

Table 5: Row-by-row check of Proposition 4.4 along the $\beta = 0.04$, $m = 16$ Leduc run at $\rho_{\mathrm{in}}^m = 0.429$. The columns s_{tr}, $s_{\mathrm{c}}(0.9117)$, $s_{\mathrm{c}}(1)$, and $s_{\mathrm{c}}(\mathrm{emp})$ are the slacks (right-hand side minus left-hand side) of the tracking row and of the center row at $\rho_J = 0.9117$, at the nonexpansive value $\rho_J = 1$, and at the measured $\rho_J^{\mathrm{emp},t}$; negative values are violations. Eight of the 24 checkpoints are shown.

H Extended Related Work

From Hannan regret to the CFR family Fictitious play (Brown, 1951) and regret matching (Hart & Mas-Colell, 2000) are the classical forerunners; Hannan regret (Hannan, 1957) is equivalent to Blackwell approachability (Abernethy et al., 2011), the framework behind recent treeplex solvers (Chakrabarti et al., 2024). CFR (Zinkevich et al., 2007) and its Monte-Carlo variant (Lanctot et al., 2009) made regret matching the inner loop of extensive-form solvers; CFR$^+$ (Tammelin, 2014), DCFR (Brown & Sandholm, 2019), and optimistic variants (Farina et al., 2019) refine the correction terms, regret matching is online mirror descent on the treeplex (Liu et al., 2022), with RM$^+$ stabilized by restarts or orthant smoothing (Farina et al., 2023), and the line traces to online convex programming (Zinkevich, 2003). The poker milestones (Bowling et al., 2015) deployed the average strategy; last-iterate convergence is the newer target. This line optimizes inner dynamics on a fixed carrier; we analyze the reference update its proximal descendants share.

Last-iterate rates and the forgetting line Last-iterate convergence is the target in zero-sum games and in constrained min-max optimization (Daskalakis & Panageas, 2019; Wei et al., 2021), in extensive-form games (Lee et al., 2021), and in deployed systems (Perolat et al., 2022). Fast last-iterate rates require forgetful algorithms (Cai et al., 2024). The same line gives reductions from average- to last-iterate convergence (Cai et al., 2025), proximal regret as a solution concept (Cai et al., 2026), and last-iterate convergence of anchored gradient (Cai & Zheng, 2026). Optimistic and adaptive methods attain fast last-iterate rates on monotone variational inequalities (Gorbunov et al., 2022; Antonakopoulos et al., 2021; Cen et al., 2021). On the optimization side the same re-centering appears as a moving anchor. Moving-anchor extragradient methods generalize fixed-anchor ones at order-optimal worst-case rates (Alcala et al., 2025b), and their stochastic variants retain an accelerated rate theory (Alcala et al., 2025a). An operator-side view identifies the anchor with a vanishing Tikhonov regularizer that unifies Halpern, extragradient, and Popov iterations (Chen, 2026). Anchored gradient descent–ascent attains an improved last-iterate $O(1/t)$ rate for the squared gradient norm on smooth convex-concave problems (Surina et al., 2026). Minimax-optimal Halpern step sizes recover the geometric Banach–Picard rate on contractive maps (Bravo et al., 2026). These analyze the anchor's motion under monotone or contractive structure. Our carrier closes the tracking error of one warm-started projected step against a moving center (Theorem 4.5). Our interior rate (Theorem 6.1) shares the forgetting intuition and applies it to the center's motion.

Proximal centers at the two ends The two ends of Section 1 are the fixed magnet of MMD (Sokota et al., 2023) and the re-centering or resets of APMD (Abe et al., 2024), RTCFR$^+$ and its adaptive variant (Meng et al., 2025; Ren et al., 2025), and Magnetic preference optimization (MPO) (Wang et al., 2025). Which member of a non-singleton equilibrium set a solver returns is itself an algorithmic choice. A tabular study of extensive-form solvers reports solver-dependent selection across the Nash polytope, with fixed-magnet and anchored methods selecting a maximal-entropy member (Leal, 2026). Perturbation under sampling sharpens their guarantees (Masaka et al.,

2025), and regularized sequence-form solvers give a complementary lens (Liu et al., 2023). Perturbing one player's payoff alone avoids the decaying perturbation that symmetric schemes require in order to return to the original equilibrium. It preserves a linear last-iterate rate to an equilibrium of the unperturbed game on bilinear saddle problems (Abe et al., 2026). A continuously updated reference also arises as a running-average anchor in normal-form play (Savcı, 2026) and as a per-round advance in two-loop alignment (Liu et al., 2026). On the treeplex carrier the moving reference is the prox center itself (Theorem 4.5). The device of moving the center also appears in proximal-point algorithms for variational inequalities, where a nonsymmetric scheme drifts its proximal center across iterations (Guo & Han, 2020). The proximal-point family bounds inexact relaxed proximal iterations under relative-error certificates or explicit correction (Pethick et al., 2025), with metric subregularity making single-point certificates closed-form (Robinson, 1981). We absorb the one-step tracking error into a coupled system and obtain the finite-time, rate-explicit statement for the interior $\beta_t \in (0, 1)$ (Theorem 4.5).