BINDING BUDGETS DESTABILIZE PRIMAL-DUAL LEARNING IN RANDOM GAMES

Anonymous authors
Paper under double-blind review

ABSTRACT

Primal-dual learners raise a Lagrange multiplier while spending exceeds a budget and lower it while under. No prior study measures how often such learners fail to converge in random games, or for which budgets. We measure both in random two-player games with per-action costs and one budget each. Outside zero-sum games, binding budgets make more runs from almost the same start separate, with a peak at intermediate budgets. A Lyapunov exponent confirms this trend for chaos. The peak is intermediate because tighter budgets bind harder but leave fewer games feasible, and infeasible games separate less often. We trace this non-convergence to delay: because the multiplier lags spending, each strategy and its multiplier overshoot and oscillate. We prove for two actions that competitive coupling enlarges the linearized oscillation, averaged over both players' independent phases. Learning the multiplier is the main cause of the added separation: learners instead projecting strategies inside the budget almost never separate on the same feasible games. Damping or a proportional term on the multiplier restores convergence on most feasible games, the latter with far less late-run spending above the budget.

1 INTRODUCTION

An advertiser that adjusts its bids over time must respect a budget or return-on-spend target (Paes Leme et al., 2024). In safe reinforcement learning, an agent must keep its spending (expected cost) below a limit (Altman, 1999; Stooke et al., 2020). A standard way to learn under a budget is primal-dual learning (Arrow et al., 1958; Flåm, 1993), which subtracts the violation (spending minus budget) times a weight from the objective and learns the weight as well. This weight is a Lagrange multiplier, the price of violating the budget: it rises while the budget is exceeded and falls while the budget is met. When primal-dual learning fails to converge, the strategies (probability distributions over actions) keep moving and the budget can be overshot repeatedly.

Without budgets, prior work uses regime diagrams: charts of where learning on random games converges, cycles, or becomes chaotic (Galla & Farmer, 2013; Sanders et al., 2018). They show that the outcome depends on the correlation of the two players' payoffs. Because no regime diagram exists for learning with budgets, we ask four questions on random two-player games with one budget per player. (1) How often does learning fail to converge, and for which budgets is non-convergence most frequent? (2) Which mechanism causes non-convergence, and which components of the learning rule does it require? (3) What does damping, which pulls the multiplier toward zero by a small fraction at every step, give up in tail violation (late-run spending above the budget)? (4) Does non-convergence persist under sampled feedback, one sampled opponent action for the payoff and one sampled own action for the cost instead of their expectations?

Our contributions are threefold. (1) **A regime diagram with budgets:** We chart the first such diagram for primal-dual learning over random games, and one budget per player changes it at almost every tested parameter combination. A budget binds when an unconstrained learner would spend more than it, and binds strongly when that learner would spend much more. Among feasible games (those whose budgets can be met), runs started from nearly the same point separate more often the more strongly the budget binds. A Lyapunov exponent measured in longer runs confirms the same trend for chaos. Tighter budgets also make more games infeasible, and these separate and turn chaotic less often. Outside zero-sum games the separating and chaotic fractions therefore both peak

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.

at intermediate budgets. (2) **An oscillator model of non-convergence:** We explain non-convergence with a model in which each strategy and its own multiplier form an oscillator, and both overshoot because the multiplier reacts to accumulated violation. Coupling is competitive when whatever raises one player's payoff difference between its two actions lowers the other's, as in every zero-sum game. For two actions we prove that both variables oscillate, that competitive coupling enlarges the oscillation on average over both oscillators' independent phases, and that under the simultaneous update only the multiplier's bounds can stop this growth. With no fitted constant, the frequency the model predicts matches the median measured frequency of cycling runs at the two larger step-size ratios (multiplier over strategy step size). In controlled comparisons, a learner that projects its strategy inside the budget instead of learning a multiplier almost never separates on the same feasible games, so the learned multiplier, rather than the budget alone, is the main cause. (3) **Two remedies and sampled feedback:** We find that damping makes most feasible games converge but raises tail violation. A proportional term on the multiplier makes as many games converge as damping, at one to two orders of magnitude lower tail violation. Sampled feedback does not restore convergence.

2 Related Work

The first of three threads studies regime diagrams of learning in random games. Sato et al. (2002) showed that reinforcement-type learning in a single two-person game can be chaotic. Galla & Farmer (2013) and Sanders et al. (2018) charted where learning over Gaussian random games converges, cycles, or turns chaotic as the payoff correlation varies, the latter also as the number of players varies. Recent work adds other axes (Chan & Galla, 2026; Garnier-Brun et al., 2024). Because none of these diagrams has a budget axis, they cannot tell whether a budget makes learning more or less stable. Without budgets our runs follow the same pattern (Section 5.1): most zero-sum runs separate (two nearby runs move apart), and cycling and then convergence prevail as the correlation rises. We add the budget axis.

The second thread proves that standard learning rules can fail to converge in specific games. A variant of multiplicative weights with a constant step can cycle or turn chaotic in congestion games (each resource gets worse the more players use it) (Palaiopanos et al., 2017). Multiplicative weights is chaotic in zero-sum games (Cheung & Piliouras, 2019; 2020). Game decompositions extend this chaos beyond zero-sum games (Cheung & Tao, 2021). In min-max problems, gradient methods can converge to spurious sets that contain no equilibrium (Daskalakis & Panageas, 2018; Hsieh et al., 2021). Optimistic and extragradient updates extrapolate the gradient by one step and restore last-iterate convergence in zero-sum games (Daskalakis & Panageas, 2019; Mertikopoulos et al., 2019). Autobidders that adapt bid multipliers to return-on-spend constraints can cycle or turn chaotic on constructed instances (Paes Leme et al., 2024; Anagnostides et al., 2026). We ask how often a learned multiplier produces such non-convergence in random games, and through which mechanism.

The third thread is primal-dual constrained learning, which goes back to Arrow et al. (1958). Convergence theory covers constrained Markov decision processes (Ding et al., 2023), constrained monotone games (Flåm, 1993; Meng & Li, 2023), and constrained Markov games (Chen et al., 2022). Holding & Lestas (2021a;b) analyze the stability of saddle-point dynamics (the coupled ascent on the primal and descent on the dual variable of a Lagrangian). Stooke et al. (2020) add proportional and derivative terms to the multiplier to reduce its oscillation in safe reinforcement learning. Primal-dual learning is a piecewise-smooth system, because the update rule switches when the multiplier reaches its lower bound of zero. Such systems have their own bifurcations when an orbit touches or crosses that boundary (Nordmark, 1991; Nusse & Yorke, 1992; Zhusubaliyev & Mosekilde, 2003; di Bernardo et al., 2008; Truong et al., 2020). We measure the visits to zero and what damping gives up in tail violation.

3 Preliminaries

A two-player game with K actions per player is given by payoff matrices $A, B \in \mathbb{R}^{K\times K}$: when player 1 plays action i and player 2 plays action j, they receive A_{ij} and B_{ij}. A mixed strategy is a probability vector over actions, $x \in \mathcal{S}_K$ for player 1 and $y \in \mathcal{S}_K$ for player 2, where the simplex $\mathcal{S}_K$ is the set of nonnegative vectors that sum to one. A pure strategy puts all its probability on one action. Expected payoffs are $x^\top Ay$ and $x^\top By$. Learning in games means that each player

repeatedly updates its mixed strategy from the payoffs it observes (Mertikopoulos, 2026). A standard rule is multiplicative weights (Arora et al., 2012). Player 1 observes the payoff vector $g = Ay$ and player 2 observes $B^\top x$. With a step size η, the rule multiplies each action's probability by the exponential of η times its payoff and renormalizes:

$$\mathrm{MW}(x, \eta g)_k \;=\; x_k e^{\eta g_k} \Big/ \textstyle\sum_j x_j e^{\eta g_j}. \tag{1}$$

Equivalently, if x is the softmax of a logit vector ℓ ($x_k \propto e^{\ell_k}$), the rule adds ηg to ℓ. If both players use it with a constant step size, this rule need not converge.

A constrained player maximizes its expected payoff subject to a budget on an expected cost: with a cost vector c and a budget d, the constraint is $x^\top c \le d$. Primal-dual learning, a standard way to learn under such a constraint, forms the Lagrangian $x^\top(Ay - \lambda c) + \lambda d$ and updates strategy and multiplier together (Arrow et al., 1958). The Lagrange multiplier $\lambda \ge 0$ is the price of violating the budget. The strategy ascends the Lagrangian by the multiplicative-weights step (1) and therefore learns on the Lagrangian payoff $Ay - \lambda c$. The multiplier moves by a step proportional to the violation $x^\top c - d$: it rises while spending exceeds the budget and falls while spending is below it. Because the multiplier never goes below zero, it never rewards spending below the budget. A strategy is the primal variable and its multiplier the dual variable.

Each run gets one of three labels: (1) converging when its strategies stop moving, (2) cycling when they keep moving but runs from nearby starts stay together, and (3) chaotic when its largest Lyapunov exponent is positive. Such convergence need not reach a Nash equilibrium, a strategy pair from which no player gains by changing its own strategy. The learning state, both strategies and their multipliers, has one Lyapunov exponent per dimension. The largest is the average exponential rate at which two runs from nearly identical starts move apart (Katok & Hasselblatt, 1995). Because runs are finite, we average this rate over part of a run, which gives the finite-time exponent.

4 Methodology

4.1 Random constrained games and how we measure them

The ensemble is the distribution of random games we study. We vary three parameters: payoff correlation ρ, budget ratio ζ, and step-size ratio τ (multiplier step size over strategy step size). A choice of the three parameters is a cell. The payoff entries are independent standard Gaussian draws, except that the two players' payoffs at the same joint action are correlated:

$$\mathrm{corr}\big(A_{ij}, B_{ij}\big) = \rho \in [-1, 1]. \tag{2}$$

$\rho = -1$ gives zero-sum games, $\rho = 1$ a common payoff, and $\rho = 0$ independent payoffs. Each player has one linear cost constraint, drawn independently of the payoffs. Player 1 draws a cost vector c_x with entries independent and uniform on $[0.5, 1.5]$, and playing action k costs $c_{x,k}$. Its budget limits its spending to ζ times the uniform-play cost $\bar{c}_x$, the mean of the cost entries in that game (what playing every action equally often costs):

$$x^\top c_x \;\le\; d_x, \qquad d_x = \zeta \cdot \bar{c}_x, \tag{3}$$

with the symmetric constraint for player 2. A smaller ζ is therefore a tighter budget, and $\zeta = \infty$ removes the budget.

Two properties of a game decide how its budget acts. One is whether the budget can be met at all. A game is feasible when each player has a strategy that meets its own budget. This holds when each budget is at least the cost of the cheapest action, $d \ge \min_k c_k$. A budget with $\zeta < 1$ does not guarantee feasibility, because costs differ between games. At $\zeta = 0.5$ only 0.004 of games are feasible for both players. The other is how strongly a budget binds. Costs are independent of payoffs, so an unconstrained learner spends the uniform-play cost on average over games. A budget with ζ near 1 binds in some games but not in others. We therefore measure each game: we run it with both players unconstrained and let $\bar{x}$ be the mean strategy of player 1 over the final 20% of that run. Player 1 has slack $m_x = \zeta - \bar{x}^\top c_x / \bar{c}_x$, the budget minus this unconstrained spending in units of the uniform-play cost. The slack m_y of player 2 is defined in the same way, and the game's slack is $m = \min(m_x, m_y)$. The budget binds when $m < 0$, as an unconstrained learner would

then exceed at least one budget, and does not bind when $m > 0$. It barely binds when m is just below zero, and binds strongly when m is below -0.5.

Both players run the primal-dual learning of Section 3 with the multiplicative-weights rule (1). With strategy step size η, each multiplier has step size $\tau\eta$. A projection $\Pi_{[0,\lambda_{\max}]}$ keeps it between 0 and a default upper bound $\lambda_{\max} = 50$. Our main update first updates the multiplier from the current violation, and the strategy step then uses the new multiplier:

$$\lambda_x^{(t+1)} = \Pi_{[0,\lambda_{\max}]}\big(\lambda_x^{(t)} + \tau\eta\,(x_t^\top c_x - d_x)\big), \qquad x_{t+1} = \mathrm{MW}\big(x_t,\ \eta\,(Ay_t - \lambda_x^{(t+1)} c_x)\big), \quad (4)$$

with the symmetric update for (y, λ_y). Section 4.2 also analyzes the simultaneous update, in which each strategy and its own multiplier step from the same time-t state.

We give each run one of three classifier labels, read from its tail, the final 20% of its T steps. A run is converged when its strategies stop moving: the tail mean of $\|x_{t+1} - x_t\|_2 + \|y_{t+1} - y_t\|_2$ is below 10^{-6}. For a run that keeps moving, we apply the separation test: we start a second copy of the run with player 1's first logit raised by 10^{-8} (a coordinate fixed before any run). Every 50 steps we take the difference Δ of the two copies' logit vectors, record $\log(\|\Delta\|_2/10^{-8})/50$, and renormalize Δ (not the multipliers) to norm 10^{-8}. We average the recorded values over the second half of the run. The run is separating when this average, the separation rate, exceeds 10^{-3} per step. Otherwise it is cycling: the copies stay together. The three labels partition the games and are those of Section 3 with one change: separating replaces chaotic, because the test is a cheap high-recall proxy and chaos is measured separately below. A game is non-converged when it is cycling or separating.

At $\zeta = 0.5$, 0.92 of the infeasible games converge to the pure strategy of their cheapest action. The multiplier reaches $\lambda_{\max}$ in most such runs, but raising $\lambda_{\max}$ to 500 leaves that converged fraction unchanged. Such a run counts as converged although its budget is never met, because the labels describe how the strategies move. Hence most runs converge at the tightest budgets. Feasibility and the dynamics therefore both shape the peak of the separating fraction at an intermediate budget. At the peak cell of the fine ζ sweep ($\rho = 0$, $\zeta = 0.75$, $\tau = 1$) only a third of the infeasible games converge that way. Raising $\lambda_{\max}$ to 500 changes the separating fraction at this peak by at most 0.01. We expected the simultaneous update to raise the separating fraction, but on the same games it moves each label fraction by at most 0.035, a negligible to slightly stabilizing effect.

We confirm chaos with the largest finite-time Lyapunov exponent of Section 3: for each game we make an additional run from the same initial logits and re-orthogonalize four tangent vectors every 20 steps. We count the exponent as positive and call the game chaotic only when the fastest-growing vector accumulates a log growth above 3 over the second half of the run, a fixed factor of e^3 at every horizon. Over those $T/2$ steps this is a per-step rate of $6/T$, and chaos always means an exponent above it. The chaotic fraction stops changing beyond $T = 3 \times 10^5$, and we quote it at that horizon unless stated. The thresholds 10^{-3} and $6/T$ are not comparable: the first is a rate on the logits alone, the second on the logits and multipliers together. The estimator stays below $6/T$ on an analytic rotation whose exponent is zero in closed form. Chaos is an extra attribute and not a fourth classifier label: a chaotic game is counted both under its classifier label and as chaotic. The separating label misses few chaotic games but covers many that are not: at the peak cell (Section 5.1) it covers 0.97 of the chaotic games, while only 0.41 of the separating games are chaotic. The exponents of the other separating games shrink toward zero as the run grows longer (Appendix D.9.10).

We study the map from a cell (ρ, ζ, τ) to the behavior of learning, with the default choices $K = 8$ actions per player, strategy step size $\eta = 0.05$, run length $T = 30{,}000$, initial logits i.i.d. $\mathcal{N}(0, 0.1^2)$, and multipliers starting at zero. The step-size ratio takes the values $\tau \in \{0.1, 1, 10\}$, which the coarse grid crosses with $\rho \in \{-1, -0.5, 0, 0.5, 1\}$ and $\zeta \in \{0.5, 0.8, 1.2\}$. The output is the fraction of G independent games per cell that carry each classifier label, with $G = 2{,}000$ everywhere except the regime diagram of Figure 2, the fine ζ sweep, and the damping sweep, which use $G = 20{,}000$. Additional longer runs give the chaotic fraction, with the number of games stated with each result.

We compare the regime diagram with and without budgets, find the cells with the largest non-converged fraction, and split each fraction into its feasible and infeasible parts, because the budget ratio also decides feasibility. For the mechanism the input is the update rule (4) and a game. One output is an oscillator model of one strategy coupled to its own multiplier, which predicts the oscillation frequency and whether the oscillation grows. The other is the list of components of (4) that a positive exponent needs. We replace one component at a time on the same games. We call a component needed when its replacement removes the chaos at the peak cell.

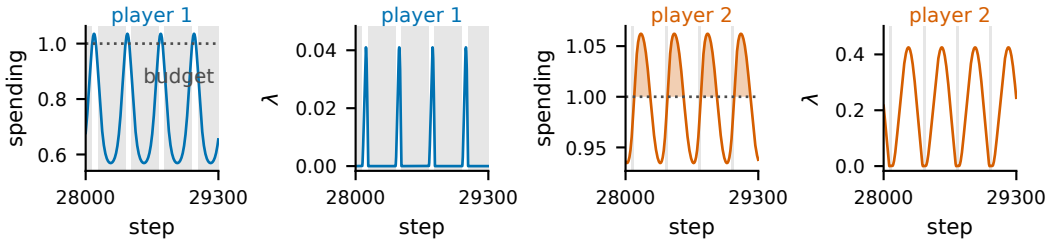


Figure 1: Each multiplier peaks exactly when its own player's spending falls back under the budget, and rests at zero in the shaded stretches, most of each period for player 1. Illustrative two-action run of (4) ($\rho = 0$, $\zeta = 1$, $\tau = 1$, steps 28,000–29,300). Separate scales; the colour fill marks steps over budget, player 1 never by more than four percent.

The three variants below are alternatives, not combined. Damping replaces the multiplier update of (4) by $\lambda_x \leftarrow \Pi_{[0,\lambda_{\max}]}(\lambda_x + \tau\eta\,(x^\top c_x - d_x - \delta\lambda_x))$. A proportional term (Stooke et al., 2020) replaces the multiplier that enters the strategy step by $\Pi_{[0,\lambda_{\max}]}(\lambda_x + k_p(x^\top c_x - d_x))$ and leaves the update of the multiplier itself unchanged. Sampled feedback replaces the exact payoff Ay in (4) by the averaged payoffs of b actions sampled from the opponent's strategy. Likewise, the exact spending $x^\top c_x$ in the multiplier update becomes the average cost of b actions sampled from x, and $b = 1$ is the main case. Tail violation is the positive part of $x^\top c_x - d_x$, averaged over the tail and the two players. We measure how it and the label fractions change with damping rate δ, proportional gain k_p, and batch size b.

4.2 The primal-dual oscillator model

The proofs give the oscillation frequency, how much competitive coupling enlarges the linearized oscillation for two actions (averaged over independent phases), and, for the two-action simultaneous update with competitive coupling, that only the multiplier's bounds can fold a growing oscillation back. The two-action proofs cover the regime where chaos is rarest: restricting each game at the peak cell to two actions drops the chaotic fraction from 0.32 to 0.056. At eight actions, our default, we measure, not prove, what chaos needs across the tested variants (Section 5.2).

A binding budget makes learning oscillate through a delayed feedback loop between a strategy and its own multiplier, which we call an oscillator. While the budget is exceeded, the multiplier rises and pushes the strategy toward cheaper actions. Once spending falls below the budget, the multiplier starts to fall. Yet it is still high while the gap below the budget accumulates, and the strategy overshoots toward cheap actions. The multiplier then falls far enough for the strategy to return to costly actions. Spending overshoots the budget again, and the loop repeats. In the two-action run of Figure 1, the multiplier of player 2 peaks exactly when its spending falls back below the budget.

Our explanation has four links: (1) a binding budget makes each oscillator oscillate, (2) coupling between the players enlarges the oscillation, (3) multipliers at or near zero fold the growing oscillation back, and (4) the result is chaotic motion with a positive exponent. We prove links (1), (2) and part of (3), each in one of the three systems named next, and measure the rest (Table 1). One is the continuous-time limit of small steps, which gives the frequency. Another is the multiplier-first update (4) that we run; its linearization gives the two energy statements, where energy measures the size of the oscillation. The third is the simultaneous update, which gives the folding statement.

The oscillator and its frequency. Consider two actions per player ($K = 2$), where only the cost difference matters. For costs (c_1, c_2) the budget is exactly met when the costly action has probability $p^* = (d - c_2)/(c_1 - c_2)$. Rescaling each player's own costs to $c = (1, 0)$ leaves that p^* unchanged and multiplies the multiplier step per unit violation by $(c_1 - c_2)^2$; we write $\tilde{\tau} = \tau\,(c_1 - c_2)^2$ for the step-size ratio after that rescaling. The frequency formula below and every measurement keep the costs as drawn, with the factor $(c_1 - c_2)^2$ entering through the variance $\mathrm{Var}_x(c)$ of the cost under the strategy; only the two-action equations of this subsection, where $c = (1, 0)$, use $\tilde{\tau}$. We write p and q for the probabilities with which the two players take their first action. In the limit of small steps, and while both multipliers stay strictly inside $[0, \lambda_{\max}]$, the update (4) becomes a differential equation, with time measured in units of $1/\eta$ steps. The strategy moves at a rate proportional to its penalized payoff difference, the payoff difference between the first action and the second minus the

multiplier. The multiplier accumulates the violation:

$$\dot{p} = p(1-p)\,\big(u(q) - \lambda_x\big), \qquad \dot{\lambda}_x = \tilde{\tau}\,\big(p - p^*\big), \tag{5}$$

and symmetrically for (q, λ_y) with its own p^* and $\tilde{\tau}$. Here $u(q)$ is the payoff difference between action 1 and action 2 against the opponent strategy q, and $p(1-p)$ is the multiplicative-weights step written in probabilities. u is linear in q with slope u', and the payoff difference $v(p)$ of player 2 is linear in p with slope v'. At the fixed point of (5) both budgets are exactly met: the two probabilities are p^* and q^*. The multipliers that hold the strategies there, $\lambda_x^* = u(q^*)$ and $\lambda_y^* = v(p^*)$, must be nonnegative. Linearizing at this point with the opponent held fixed turns each oscillator into a center, a fixed point around which the linearized motion is a pure rotation. The lemma below gives the oscillator's frequency in closed form at every K, with Δ a small deviation from the fixed point.

Lemma 1 (Frequency formula, coefficients frozen at the fixed point)**.** *For K actions and cost vector c, hold the payoff-cost covariance and the variance of the cost under the strategy fixed at their values at a fixed point where both budgets are met. Linearizing the multiplier-driven motion of spending there gives $\frac{\mathrm{d}}{\mathrm{d}t}(x^\top c) = -\eta\,\mathrm{Var}_x(c)\,\Delta\lambda$ and $\frac{\mathrm{d}}{\mathrm{d}t}\lambda = \tau\eta\,\Delta(x^\top c)$ (time in steps), a harmonic oscillator whose per-step angular frequency is*

$$\omega \;=\; \eta\sqrt{\tau\ \mathrm{Var}_x(c)}, \tag{6}$$

with $\mathrm{Var}_x(c)$ the variance of the costs under the strategy. At $K = 2$, $c = (1, 0)$, this is $\eta\sqrt{p^*(1-p^*)\,\tau}$.

Energy: conserved alone, injected by coupling. For the linearized multiplier-first update the energy is $Q = \tilde{\tau}\alpha\,\delta\ell^2 + \tilde{\tau}\eta\alpha\,\delta\ell\,\delta\lambda + \delta\lambda^2$ in the deviations of the logit of p and of the multiplier from the fixed point, with $\alpha = p^*(1-p^*)$. Under that update, the linearized oscillator whose opponent is held fixed conserves this energy, and its orbits therefore stay on ellipses (Proposition 1). Coupling is competitive when $u'v' < 0$, so that whatever raises the payoff difference of one player lowers that of the other. In the ensemble u' and v' are Gaussian with correlation ρ: zero-sum games are always competitive, common-payoff games never are, and at $\rho = 0$ half of the two-action games are. Under competitive coupling each oscillation drives the other in the direction that enlarges it, and the energy grows on average over a period when the two phases are independent (Proposition 2). The factor α scales this growth and the predicted frequency: both are largest at $p^* = 1/2$ and vanish as the budget tightens ($p^* \to 0$). This growth is a property of the discrete update, since in the continuous-time limit competitive coupling only splits the two frequencies (Corollary 1). Chaos needs a positive Lyapunov exponent and bounded orbits, so that the region they fill is repeatedly stretched and folded back onto itself (Strogatz, 2024). Competitive coupling is therefore the candidate for stretching.

Folding happens only at a bound (simultaneous update). Folding could come from saturation of the strategy (a probability cannot exceed one), coupling to the other player, or the multiplier's

Table 1: What is proven or measured, and for which system. Continuous-time limit: two actions unless stated. Multiplier-first update: the update we run; the two energy rows concern its linearization. Simultaneous update: both variables step from the same state.

Statement	System, scope	Status
Frequency formula (Lemma 1)	continuous-time limit, any K	proven; agrees with measurement at $\tau \geq 1$, not at $\tau = 0.1$ (Section 5.2)
Oscillation at the fixed point (Cor. 1)	continuous-time limit, $K = 2$	proven
Isolated oscillator conserves energy (Prop. 1)	multiplier-first update, any K	proven; verified symbolically
Competitive coupling injects energy (Prop. 2)	multiplier-first update, $K = 2$	proven, for independent phases
No folding away from the bounds (Prop. 3)	simultaneous update, $K = 2$	proven (competitive coupling); multiplier-first update: the drop transfers, the growth does not
Isolated oscillator at zero has a non-positive top exponent (Remark 7)	simultaneous update, $K = 2$	proven under two assumptions
Two coupled oscillators at zero: positive exponents	multiplier-first update, $K = 2$	measured
Leading Lyapunov vector on low-probability actions (Table A6), growth retained in the oscillator directions (Table A22)	multiplier-first update, $K = 8$	measured

bounds. The argument uses the simultaneous update, because there an isolated linearized oscillator already spirals outward at every step (Remark 2). In the two-action simultaneous update saturation and coupling cannot fold when the coupling is competitive. There the continuous motion of an isolated oscillator conserves an energy H. Every step off the fixed point increases H, and a weighted version of H does the same for any two competitively coupled oscillators (Proposition 3). Every such orbit other than the fixed point therefore leaves every compact part of the region strictly inside the multipliers' bounds, so any folding must happen at a bound. The bound at zero does fold. While a multiplier is held at zero, the half-period that would take it below zero becomes a rest. The oscillation leaves zero with an energy set by the bound at zero, whatever energy it arrived with. Large and small oscillations map onto similar ones. Under the two properties of Remark 7, an isolated oscillator at zero has a non-positive top exponent; it visits zero once per revolution, with closed-form amplitude and period within 10% of simulation and no chaos in any of the 54 simulated parameter settings.

Multiplier-first update, two actions (measured). In simulation the isolated nonlinear oscillator of this update shows no energy growth over a period. The energy drop at a visit to zero transfers from the simultaneous update in 0.96 of visits, but the energy growth away from zero holds for only 0.59 of competitively coupled oscillators, below the three-quarters criterion fixed in advance (Remark 4). With two actions, positive exponents appear only when two oscillators that visit zero drive each other.

Eight actions (measured). With eight actions at $\rho = 0$ and $\tau = 1$, the separating fraction among feasible games rises steadily as the budget tightens (Section 5.1), unlike the two-action frequency and energy growth. In eight-action runs perturbation growth is retained in the oscillator directions, the multipliers and cost-gradient directions. At the peak cell 93% of the eight-action games reach zero within 3×10^5 steps (Appendix D.10.1), yet raising the lower bound or replacing it by a softplus reparameterization leaves the separating fraction near that of the unchanged learner (Table 2, Section 5.2), so at eight actions chaos does not need the bound at zero itself.

5 Experiments

All fractions in this section are fractions of games. Per-cell fractions at $G = 2{,}000$ carry $3\sigma \leq \pm 0.034$ (binomial over the game draws), and the 400-game chaotic fractions at most ± 0.075. The separating fraction is our main measure, and the chaotic fraction, judged at $K \geq 8$ by criteria fixed before the runs, confirms its main results. Appendix A holds the proofs, Appendices B and C the setup and the measurements, and Appendix D every experiment summarized here.

5.1 Separating games peak at an intermediate budget ratio

Each row of the regime diagram, one row per payoff correlation, changes once each player has a budget. In Figure 2 every row of every panel changes shade along ζ, and 43 of the 45 coarse-grid cells differ from their no-budget value beyond 3σ. Every row except the zero-sum row ($\rho = -1$) develops a peak of separating games at intermediate ζ, because at the zero-sum row most runs

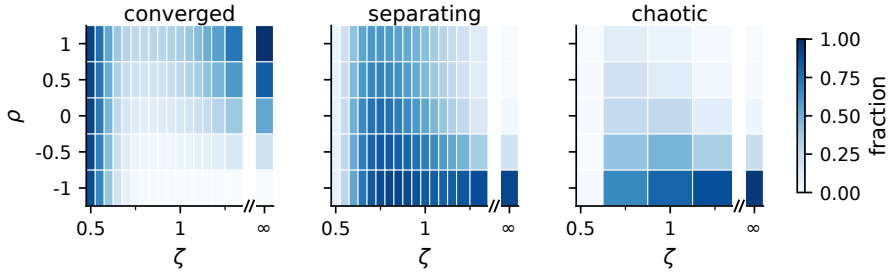


Figure 2: One budget per player changes every row of the regime diagram. Converged and separating fractions and the chaotic fraction over (ρ, ζ) at $\tau = 1$ and $K = 8$; the cycling fraction is what the first two leave, and the strip marked ∞ has no budget. At the tightest budgets most games cannot be met, and those runs converge at a permanent violation. Classifier panels: ζ from 0.5 to 1.3, $20{,}000$ games per cell, $T = 30{,}000$. Chaotic panel: 400 games per cell, $T = 3 \times 10^5$; its cells at $\zeta = 0.25$ and 1.5 lie outside the shared axis.

separate even without budgets. There a tight budget moves most games to a pure strategy, where they converge. The separating fraction returns to about its no-budget value as the budget loosens.

At $\rho = 0$ the separating fraction peaks at an intermediate budget at every step-size ratio, and stays near 0.03 without budgets (left panel of Figure 3). At $\tau = 1$ the maximum is 0.76 at $\zeta = 0.75$, and the other step-size ratios peak one to two grid steps away. The peak comes from two opposite trends (right panel). As ζ rises from 0.5 to 0.95 the feasible fraction rises from near zero to one. Over the whole sweep to $\zeta = 1.3$ the separating fraction among feasible games falls from 0.84 to 0.15. Their product, the feasible games' contribution, peaks in between, and near the peak infeasible games separate less often than feasible ones. This peak therefore lies where most games have just become feasible. From each of five held-out cost distributions alone, we compute the budget ratio at which the feasible fraction reaches its value at our peak. This ratio predicts the held-out peak within one or two grid steps (Appendix D.9.20). The peak persists in the robustness variants of Appendix D.2, which change the payoffs, the costs, the game size, or the update rule.

The chaotic fraction confirms this peak: it peaks within one grid step of the separating-fraction peak, at about one third of the games (black squares, left panel of Figure 3). In $2{,}000$ games at the peak cell ($\rho = 0$, $\zeta = 0.75$, $\tau = 1$) run for $T = 3 \times 10^5$, the chaotic fraction is 0.375 among feasible games and 0.036 among infeasible ones, against 0.044 without budgets. There the chaotic fraction stays at about one third of all games through 10^7 steps.

Among feasible games whose no-budget run converged, binned by slack m, 0.79 separate where the budget binds strongly, 0.59 where it barely binds, and 0.03 to 0.08 once it does not bind. The chaotic fraction follows the same order, and the order also holds within a single budget ratio (Appendix D.3).

5.2 The learned multiplier drives separating runs and chaos

The frequency formula predicts the oscillation frequency of cycling runs at the two larger step-size ratios (Figure 4). We test this prediction, $\omega = \eta\sqrt{\tau \operatorname{Var}_x(c)}$, on $5{,}186$ cycling player–game pairs at $\rho = 0$ over the three step-size ratios. We keep the $2{,}801$ that pass an inclusion rule fixed before the fit: a multiplier away from both its bounds, a cost variance above 10^{-6}, and one dominant spectral peak. Their median ratio of measured to predicted frequency is 1.03 at $\tau = 1$ and 1.02 at $\tau = 10$, against 1.53 at $\tau = 0.1$. The pooled log-log slope is 0.82, and 0.92 for pairs predicted well above the resolution limit, the lowest frequency a run can resolve. At $\tau = 0.1$ the oscillations lie near the resolution limit, so the frequency test fixed before the runs failed there.

Learning the multiplier accounts for most of the separating fraction, as two projection learners show. They keep the strategy feasible by Euclidean or Kullback–Leibler (KL) projection onto the budget set instead of a learned multiplier, with the same step size and draws. On the same feasible games both learners almost never separate, and the KL learner, which keeps multiplicative weights, is also rarely chaotic (Table 2).

Across the tested variants, the peak of the separating fraction needs the multiplier's integral of past violations but not the exact bound at zero (Table 2), and so does the chaotic fraction where we measure it (Appendix D.5). Raising the lower bound to 1, still well below the median active multiplier of 4.3, lowers the separating fraction at the peak cell from 0.77 to 0.61.

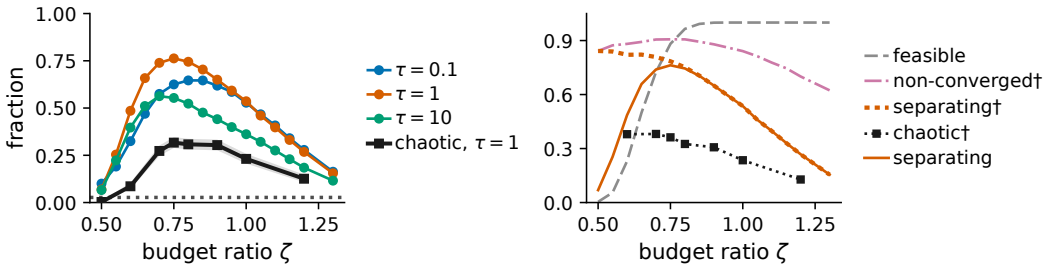


Figure 3: The total separating fraction peaks where the feasible fraction rises while the separating fraction among feasible games falls ($\rho = 0$); a dagger marks fractions among feasible games. Left: separating per step-size ratio; chaotic at $\tau = 1$ (squares) and no budgets (dotted). Right, at $\tau = 1$: the feasibility split.

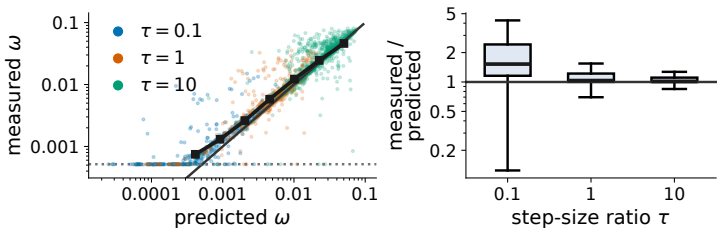


Figure 4: The frequency formula predicts the measured frequency with no fitted constant at the two larger step-size ratios ($\rho = 0$, included cycling pairs). Left: pairs by τ, binned medians over the two larger ratios (squares), equality line (solid), resolution limit (dotted). Right: measured over predicted per τ (box: quartiles; whiskers: 5th to 95th percentiles; line: median).

Table 2: Which changes remove the peak, meaning that the separating or chaotic fraction there falls to its no-budget value. The penalty variant reads only the current violation, the sliding window only recent ones. Unless stated, entries are at the peak cell and over all games, converged fractions and tail violations among feasible games; classifier fractions at $T = 3 \times 10^4$, chaotic at 3×10^5.

Change	Measured against the unchanged learner on the same draws	Peak
penalty variant; sliding window	separating ≤ 0.023 (penalty variant) and 0.02–0.03 (sliding window), against 0.77 unchanged	removed
raised lower bound	separating 0.71–0.77 at $\lambda_{\min} \leq 0.3$, 0.79 with a softplus reparameterization, against 0.77 unchanged	kept
damping	separating 0.03 at $\delta = 1$; chaotic 0.20 to 0.09 at $\delta = 0.003$ to 0.03 against 0.29 unchanged, these three at $T = 1.2 \times 10^5$	removed
proportional term ($k_p = 30$)	converged 0.76 at tail violation 0.0015, against 0.76 at 0.17 for damping with $\delta = 0.3$; chaotic 0.012 against 0.321	removed
projection learners	separating ≤ 0.032 against 0.28–0.82 among feasible games over ζ 0.6 to 1.2; KL chaotic 0.012 against 0.321	removed

5.3 A PROPORTIONAL TERM RESTORES CONVERGENCE; SAMPLED FEEDBACK DOES NOT

A proportional term makes as many games converge as damping with a large δ, at one to two orders of magnitude lower tail violation (Table 2). In longer runs at the peak cell ($T = 3 \times 10^5$) success means converging with tail violation at most 0.005 and exploitability (the payoff players could still gain within budget) at most 0.02. The proportional term succeeds on 0.80 of feasible games. Damping with $\delta = 0.3$ succeeds on 0.03, converging as many games but leaving tail violation 0.18.

At batch one sampled feedback leaves the converged fraction and tail violation in place. With common random numbers no run is separating or chaotic at batch one: copies with shared draws keep their gap until a sampled action differs. Nearby runs separate again by batch 4,096.

6 CONCLUSION

In random two-player games away from zero-sum, a binding budget that can be met makes primal-dual runs separate and turn chaotic far more often than without budgets. The peak is intermediate because tighter budgets bind harder but leave fewer games feasible (Figure 3). The learned multiplier is the main cause: projection learners almost never separate on the same feasible games. For two actions we prove that competitive coupling enlarges the linearized oscillation on average over independent phases. Under the two-action simultaneous update, we also prove that a competitively coupled oscillation can stop growing only at a multiplier bound. At eight actions, where these proofs do not reach, the tested variants show that chaos needs the multiplier's integral of past violations but not the exact bound at zero (Appendix D.5). Extending the proofs to eight actions is the next step. In practice, a proportional term succeeds on more feasible games than damping, where success means converging with small tail violation and small exploitability.

AI USE STATEMENT

The authors set the research question, the hypotheses, the experimental standards, and every decision in this paper. Under their direction, generative AI tools took part in developing the theoretical model and its propositions, formulating and drafting proofs, designing and implementing experiments,

interpreting results, finding and formatting references, and drafting and editing code, figures, and text. No tool generated data. The provided scripts check the symbolic identities of the proofs, and the provided programs regenerate every reported number. The authors take full responsibility for the content.

Reproducibility statement

Section 4 and Appendices B and C specify the ensemble, the update rules, the thresholds, and the estimator. Every number derives deterministically from recorded integer seeds. The supplementary material holds the complete code, the per-cell results and per-game records, and the verification scripts that check the symbolic identities of Appendix A.

References

Eitan Altman. *Constrained Markov Decision Processes*. Stochastic Modeling Series. Chapman & Hall/CRC, 1999.

Ioannis Anagnostides, Ian Gemp, Georgios Piliouras, and Kelly Spendlove. Chaos in autobidding auctions, 2026. arXiv:2602.09118.

Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(6):121–164, 2012.

Kenneth J. Arrow, Leonid Hurwicz, and Hirofumi Uzawa. *Studies in Linear and Non-Linear Programming*. Stanford University Press, 1958.

Desmond Chan and Tobias Galla. Complex dynamics in the Sherrington–Kirkpatrick game, 2026. arXiv:2607.02422.

Ziyi Chen, Shaocong Ma, and Yi Zhou. Finding correlated equilibrium of constrained Markov game: A primal-dual approach. In *Advances in Neural Information Processing Systems*, 2022.

Yun Kuen Cheung and Georgios Piliouras. Vortices instead of equilibria in minmax optimization: Chaos and butterfly effects of online learning in zero-sum games. In *Proceedings of the 32nd Conference on Learning Theory*, 2019.

Yun Kuen Cheung and Georgios Piliouras. Chaos, extremism and optimism: Volume analysis of learning in games. In *Advances in Neural Information Processing Systems*, volume 33, 2020.

Yun Kuen Cheung and Yixin Tao. Chaos of learning beyond zero-sum and coordination via game decompositions. In *International Conference on Learning Representations*, 2021.

Constantinos Daskalakis and Ioannis Panageas. The limit points of (optimistic) gradient descent in min-max optimization. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Constantinos Daskalakis and Ioannis Panageas. Last-iterate convergence: Zero-sum games and constrained min-max optimization. In *Innovations in Theoretical Computer Science Conference*, 2019.

Mario di Bernardo, Christopher J. Budd, Alan R. Champneys, and Piotr Kowalczyk. *Piecewise-smooth Dynamical Systems: Theory and Applications*. Springer, 2008.

Dongsheng Ding, Chen-Yu Wei, Kaiqing Zhang, and Alejandro Ribeiro. Last-iterate convergent policy gradient primal-dual methods for constrained MDPs. In *Advances in Neural Information Processing Systems*, 2023.

Sjur Didrik Flåm. Paths to constrained Nash equilibria. *Applied Mathematics and Optimization*, 27: 275–289, 1993.

Tobias Galla and J. Doyne Farmer. Complex dynamics in learning complicated games. *Proceedings of the National Academy of Sciences*, 110(4):1232–1236, 2013.

Jérôme Garnier-Brun, Michael Benzaquen, and Jean-Philippe Bouchaud. Unlearnable games and "satisficing" decisions: A simple model for a complex world. *Physical Review X*, 14:021039, 2024.

Thomas Holding and Ioannis Lestas. Stability and instability in saddle point dynamics—Part I. *IEEE Transactions on Automatic Control*, 66(7):2933–2944, 2021a. doi: 10.1109/TAC.2020.3019375.

Thomas Holding and Ioannis Lestas. Stability and instability in saddle point dynamics—Part II: The subgradient method. *IEEE Transactions on Automatic Control*, 66(7):2945–2960, 2021b. doi: 10.1109/TAC.2020.3019381.

Ya-Ping Hsieh, Panayotis Mertikopoulos, and Volkan Cevher. The limits of min-max optimization algorithms: Convergence to spurious non-critical sets. In *Proceedings of the 38th International Conference on Machine Learning*, 2021.

Anatole Katok and Boris Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*, volume 54 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1995.

Min Meng and Xiuxian Li. Linear last-iterate convergence for continuous games with coupled inequality constraints. In *62nd IEEE Conference on Decision and Control*, pp. 1076–1081, 2023.

Panayotis Mertikopoulos. Regret, equilibrium, and learning in games: A guided tour, 2026. arXiv:2608.09389.

Panayotis Mertikopoulos, Bruno Lecouat, Houssam Zenati, Chuan-Sheng Foo, Vijay Chandrasekhar, and Georgios Piliouras. Optimistic mirror descent in saddle-point problems: Going the extra (gradient) mile. In *International Conference on Learning Representations*, 2019.

Arne B. Nordmark. Non-periodic motion caused by grazing incidence in an impact oscillator. *Journal of Sound and Vibration*, 145(2):279–297, 1991.

Helena E. Nusse and James A. Yorke. Border-collision bifurcations including "period two to period three" for piecewise smooth systems. *Physica D: Nonlinear Phenomena*, 57:39–57, 1992.

Renato Paes Leme, Georgios Piliouras, Jon Schneider, Kelly Spendlove, and Song Zuo. Complex dynamics in autobidding systems. In *Proceedings of the 25th ACM Conference on Economics and Computation*, 2024.

Gerasimos Palaiopanos, Ioannis Panageas, and Georgios Piliouras. Multiplicative weights update with constant step-size in congestion games: Convergence, limit cycles and chaos. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

James B. T. Sanders, J. Doyne Farmer, and Tobias Galla. The prevalence of chaotic dynamics in games with many players. *Scientific Reports*, 8:4902, 2018.

Yuzuru Sato, Eizo Akiyama, and J. Doyne Farmer. Chaos in learning a simple two-person game. *Proceedings of the National Academy of Sciences*, 99(7):4748–4751, 2002.

Adam Stooke, Joshua Achiam, and Pieter Abbeel. Responsive safety in reinforcement learning by PID Lagrangian methods. In *International Conference on Machine Learning*, 2020.

Steven H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. CRC Press, third edition, 2024.

Tri Quoc Truong, Tadashi Tsubone, Munehisa Sekikawa, and Naohiko Inaba. Border-collision bifurcations and Arnol'd tongues in two coupled piecewise-constant oscillators. *Physica D: Nonlinear Phenomena*, 401:132148, 2020.

Zhanybai T. Zhusubaliyev and Erik Mosekilde. *Bifurcations and Chaos in Piecewise-Smooth Dynamical Systems*. World Scientific, 2003.

A PROOFS

Table 1 in the body links the claims of Section 4.2 to the results below, with the system each is proven for. Appendix A.1 fixes the setting and the notation, Appendix A.2 proves the results on the linearized continuous-time limit, and Appendix A.3 proves the frequency formula. Appendix A.4 proves the two energy propositions for the multiplier-first update, and Appendix A.5 proves the proposition of no folding away from the multiplier's bounds for the simultaneous update. Appendix A.6 proves the criterion under which a multiplier reaches zero. Appendix A.7 is a supporting derivation of the periodic orbit.

A.1 SETTING AND NOTATION

Throughout this appendix the game has two actions per player, $K = 2$, and player 1's cost vector is $c = (1, 0)$. Write $x = (p, 1-p)$ and $y = (q, 1-q)$ for the two strategies, so that p and q are the probabilities of the first actions and player 1's spending is $x^\top c = p$. Since the uniform-play cost is $\bar{c} = 1/2$, the budget (3) is $p \le \zeta/2$. Write $u(q)$ for the payoff difference between player 1's action 1 and action 2 under the opponent strategy y. u is affine in q with slope $u' = A_{11} - A_{12} - A_{21} + A_{22}$, and $v(p)$, v' are the same objects for player 2. Competitive coupling means $u'v' < 0$, and coordination coupling means $u'v' > 0$. Write $\ell = \log(p/(1-p))$ for the logit of p and $\sigma(\ell) = 1/(1+e^{-\ell})$ for its inverse, so $p = \sigma(\ell)$ and $\sigma' = p(1-p)$. Write $p^* = \zeta/2$ for the probability at which the budget is exactly met, with $\alpha = p^*(1-p^*)$ and $\beta = q^*(1-q^*)$. An oscillator is one player's strategy with its own multiplier, and an isolated oscillator is an oscillator with the payoff difference held constant. Q, Q_w, H, and H_w are the energies that the statements below introduce. A δ in front of a symbol denotes its deviation from the fixed point.

Assumption 1 (Both budgets met with equality)**.** *Both budgets are exactly met at the fixed point, where $p^* = q^* = \zeta/2 \in (0, 1)$. Multiplier values that hold it there are nonnegative: $\lambda_x^* = u(q^*) \ge 0$ and $\lambda_y^* = v(p^*) \ge 0$.*

Assumption 2 (Interior region)**.** *Unless a statement says otherwise, both multipliers are strictly inside the interval $[0, \lambda_{\max}]$, which leaves the projection in (4) inactive.*

Two updates with the same continuous-time limit appear. The multiplier-first update is (4), in which each multiplier updates first and its strategy uses the new value. The simultaneous update is the Euler step in which every coordinate steps from the state at time t. In logit coordinates the isolated oscillator is $F(\ell, \lambda) = (\ell, \lambda) + \eta f(\ell, \lambda)$ with

$$f(\ell, \lambda) = \big(u - \lambda,\ \tau(\sigma(\ell) - \zeta/2)\big), \tag{7}$$

where u is the constant payoff difference. Its fixed point is $\sigma(\ell^*) = p^*$, $\lambda^* = u$. We write $\omega_0 = \sqrt{\alpha\tau}$ for the frequency of the linearized continuous-time limit. The continuous-time limit of both updates is the replicator flow of the strategies together with an integrator for the multipliers. Appendix A.2 derives it. Figure A1 shows the statements of this appendix on one oscillator. On the left ($\eta = 0.2$, 260 steps) the linearized multiplier-first update stays on an ellipse of the invariant Q of Proposition 1, while the simultaneous update from the same start spirals outward (Remark 2). In the middle ($u' = 1$, $v' = -1$, $\eta = 0.05$, 2,000 independent phases on the unit ellipses) the mean of Q_w grows monotonically under competitive coupling and stays constant without coupling. The mean change of Q_w over the first step is within 2% of (51) (Proposition 2). On the right ($u = 0.5$, $\eta = 0.2$, $\lambda_{\max} = 50$) a start near the fixed point and a start far from it end on the same orbit, which visits zero once per revolution (Proposition 4). The visit to zero maps a large and a small oscillation onto the same orbit.

A.2 THE TWO-ACTION CONTINUOUS-TIME LIMIT AND ITS LINEARIZATION

Both the frequency formula and the argument for no folding away from the multiplier's bounds start from the linearized continuous-time limit at the fixed point.

Lemma 2 (The two-action continuous-time limit)**.** *Under Assumption 2 and in the continuous-time limit with rescaled time $t' = \eta t$ that absorbs the step size into the time unit, the update (4) for $K = 2$ and $c = (1, 0)$ becomes*

$$\dot{p} = p(1-p)\big(u(q) - \lambda_x\big), \qquad \dot{\lambda}_x = \tau\,(p - \zeta/2), \tag{8}$$

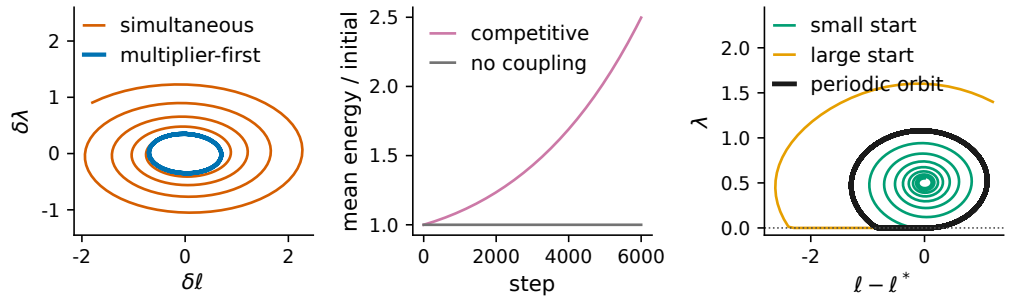


Figure A1: Phase planes of the two-action oscillator ($\tau = 1$, $p^* = 1/2$). Left: linearized isolated oscillator, multiplier-first update (ellipse) and simultaneous update (outward spiral) from one start. Middle: mean weighted invariant Q_w of two competitively coupled linearized oscillators under the multiplier-first update. Right: nonlinear isolated oscillator under the simultaneous update with the bound at zero; two starts end on one periodic orbit (black).

and symmetrically $\dot{q} = q(1-q)\big(v(p) - \lambda_y\big)$, $\dot{\lambda}_y = \tau\,(q - \zeta/2)$ *for player* 2. *Under Assumption 1 its fixed point is* $p^* = q^* = \zeta/2$, $\lambda_x^* = u(q^*)$, $\lambda_y^* = v(p^*)$.

Proof. Under Assumption 2 the projection is inactive, and the strategy step and the multiplier step of (4) are then smooth maps of the state. The vector $g = (g_1, g_2) = (u(q) - \lambda_x, 0)$ differs from the Lagrangian payoffs by a common shift that the update ignores. By Lemma 4 of Appendix A.3, whose proof does not use this lemma, the multiplicative-weights update (1) on g has the continuous-time limit

$$\dot{x}_k = \eta\, x_k\big(g_k - x^\top g\big), \qquad k = 1, 2.$$

For $K = 2$ write $x = (p, 1-p)$. Then

$$\begin{aligned} g_1 - x^\top g &= g_1 - p\, g_1 - (1-p)\, g_2 \\ &= (1-p)\,(g_1 - g_2) \\ &= (1-p)\,\big(u(q) - \lambda_x\big), \end{aligned} \tag{9}$$

and the first coordinate of the replicator flow becomes

$$\begin{aligned} \dot{p} &= \eta\, p\,\big(g_1 - x^\top g\big) \\ &= \eta\, p(1-p)\,\big(u(q) - \lambda_x\big). \end{aligned} \tag{10}$$

Inside the interval the multiplier update of (4) is

$$\begin{aligned} \lambda_x^{(t+1)} - \lambda_x^{(t)} &= \tau\eta\,\big(x_t^\top c - d_x\big) \\ &= \tau\eta\,\big(p_t - \zeta/2\big), \end{aligned} \tag{11}$$

because $x^\top c = p$ for $c = (1, 0)$ and $d_x = \zeta/2$ by (3). Dividing both increments by η and absorbing the step size into the time unit gives

$$\dot{p} = p(1-p)\big(u(q) - \lambda_x\big), \qquad \dot{\lambda}_x = \tau\,(p - \zeta/2),$$

which is (8). Player 2's computation is identical with (q, λ_y, v) in place of (p, λ_x, u). At a fixed point all four right-hand sides vanish. Both dual equations give

$$\tau\,(p^* - \zeta/2) = 0, \qquad \tau\,(q^* - \zeta/2) = 0 \quad \Longrightarrow \quad p^* = q^* = \zeta/2,$$

and since $p^*(1-p^*) > 0$ and $q^*(1-q^*) > 0$ for $\zeta/2 \in (0, 1)$, the two primal equations give

$$p^*(1-p^*)\big(u(q^*) - \lambda_x^*\big) = 0 \implies \lambda_x^* = u(q^*), \qquad q^*(1-q^*)\big(v(p^*) - \lambda_y^*\big) = 0 \implies \lambda_y^* = v(p^*).$$

Both multipliers are nonnegative exactly when $u(q^*) \geq 0$ and $v(p^*) \geq 0$, which is Assumption 1. □

Lemma 3 (Characteristic polynomial). *Let* $\alpha = p^*(1-p^*)$ *and* $\beta = q^*(1-q^*)$. *At the fixed point of Lemma 2 the* 4×4 *Jacobian of* (8) *has characteristic polynomial in the eigenvalue variable* s

$$s^4 + (\alpha + \beta)\,\tau\, s^2 + \alpha\beta\tau^2 - \alpha\beta\, u' v'\, s^2. \tag{12}$$

Proof. Order the state as $(p, \lambda_x, q, \lambda_y)$ and differentiate the right-hand sides of (8) at the fixed point of Lemma 2. In the first row the diagonal entry vanishes because the parenthesis vanishes at the fixed point,

$$\begin{aligned}\frac{\partial \dot{p}}{\partial p}\Big|_* &= (1-2p^*)\big(u(q^*) - \lambda_x^*\big) + p^*(1-p^*)\cdot 0 \\ &= (1-2p^*)\cdot 0 \\ &= 0,\end{aligned} \tag{13}$$

and the remaining entries of the first two rows are

$$\begin{aligned}&\frac{\partial \dot{p}}{\partial \lambda_x}\Big|_* = -\,p^*(1-p^*) = -\alpha, && \frac{\partial \dot{p}}{\partial q}\Big|_* = p^*(1-p^*)\,u' = \alpha u', \\ &\frac{\partial \dot{\lambda}_x}{\partial p}\Big|_* = \tau, && \frac{\partial \dot{\lambda}_x}{\partial \lambda_x}\Big|_* = \frac{\partial \dot{\lambda}_x}{\partial q}\Big|_* = \frac{\partial \dot{\lambda}_x}{\partial \lambda_y}\Big|_* = 0.\end{aligned} \tag{14}$$

For player 2 the same computation gives $\partial \dot{q}/\partial \lambda_y = -\beta$, $\partial \dot{q}/\partial p = \beta v'$, $\partial \dot{\lambda}_y/\partial q = \tau$, and zeros elsewhere, so the Jacobian is

$$J = \begin{pmatrix} 0 & -\alpha & \alpha u' & 0 \\ \tau & 0 & 0 & 0 \\ \beta v' & 0 & 0 & -\beta \\ 0 & 0 & \tau & 0 \end{pmatrix}, \qquad sI - J = \begin{pmatrix} s & \alpha & -\alpha u' & 0 \\ -\tau & s & 0 & 0 \\ -\beta v' & 0 & s & \beta \\ 0 & 0 & -\tau & s \end{pmatrix}. \tag{15}$$

Expand $\det(sI - J)$ along the second row, whose only nonzero entries are $(sI - J)_{21} = -\tau$ and $(sI - J)_{22} = s$:

$$\det(sI - J) = (-1)^{2+1}\,(-\tau)\,M_{21} + (-1)^{2+2}\,s\,M_{22},$$

with M_{21} and M_{22} the corresponding minors. The first minor expands along its first column,

$$\begin{aligned}(-1)^{2+1}(-\tau)\,M_{21} &= \tau\,\det\begin{pmatrix} \alpha & -\alpha u' & 0 \\ 0 & s & \beta \\ 0 & -\tau & s \end{pmatrix} \\ &= \tau\,\alpha\,\det\begin{pmatrix} s & \beta \\ -\tau & s \end{pmatrix} \\ &= \tau\,\alpha\,(s^2 + \beta\tau).\end{aligned} \tag{16}$$

The second minor expands along its first row,

$$\begin{aligned}s\,M_{22} &= s\,\det\begin{pmatrix} s & -\alpha u' & 0 \\ -\beta v' & s & \beta \\ 0 & -\tau & s \end{pmatrix} \\ &= s\Big(s\,\det\begin{pmatrix} s & \beta \\ -\tau & s \end{pmatrix} + \alpha u'\,\det\begin{pmatrix} -\beta v' & \beta \\ 0 & s \end{pmatrix}\Big) \\ &= s\Big(s\,(s^2 + \beta\tau) + \alpha u'\,(-\beta v' s)\Big) \\ &= s\Big(s\,(s^2 + \beta\tau) - \alpha\beta u' v'\,s\Big).\end{aligned} \tag{17}$$

Adding (16) and (17),

$$\begin{aligned}\det(sI - J) &= \tau\alpha\,(s^2 + \beta\tau) + s^2(s^2 + \beta\tau) - \alpha\beta\,u'v'\,s^2 \\ &= (s^2 + \alpha\tau)(s^2 + \beta\tau) \;-\; \alpha\beta\,u'v'\,s^2 \\ &= s^4 + (\alpha + \beta)\,\tau\,s^2 + \alpha\beta\tau^2 - \alpha\beta\,u'v'\,s^2,\end{aligned} \tag{18}$$

which is (12). □

Corollary 1 (Centers and their coupling)**.** *For $u'v' = 0$ the eigenvalues of J are two imaginary pairs $\pm i\sqrt{\alpha\tau}$ and $\pm i\sqrt{\beta\tau}$: each oscillator is a center with continuous-time frequency $\sqrt{p^*(1-p^*)\,\tau}$. Competitive coupling $u'v' < 0$ splits the two imaginary pairs, giving the linearized continuous-time limit two distinct frequencies. Coordination coupling $u'v' > 0$ makes the fixed point hyperbolic at $\alpha = \beta$. At $\alpha = \beta$ of the continuous-time limit the eigenvalues are exactly $\pm\frac{\alpha g}{2} \pm i\sqrt{\alpha\tau - \alpha^2 g^2/4}$ with $g = \sqrt{u'v'}$, off the imaginary axis for every positive coupling and fully real for $\alpha g^2 \geq 4\tau$.*

Proof. At $u'v' = 0$ the last term of (18) vanishes and

$$\begin{aligned}\det(sI - J) &= (s^2 + \alpha\tau)(s^2 + \beta\tau) \\ &= \big(s - i\sqrt{\alpha\tau}\big)\big(s + i\sqrt{\alpha\tau}\big)\big(s - i\sqrt{\beta\tau}\big)\big(s + i\sqrt{\beta\tau}\big),\end{aligned} \tag{19}$$

so the eigenvalues are the two imaginary pairs $\pm i\sqrt{\alpha\tau}$ and $\pm i\sqrt{\beta\tau}$. With $\alpha = p^*(1-p^*)$ the frequency of player 1's oscillator is $\sqrt{p^*(1-p^*)\,\tau}$. For $u'v' \neq 0$ substitute $w = s^2$ in (18):

$$\begin{aligned}0 &= (w + \alpha\tau)(w + \beta\tau) - \alpha\beta\, u'v'\, w \\ &= w^2 + \big((\alpha + \beta)\tau - \alpha\beta\, u'v'\big)\, w + \alpha\beta\tau^2,\end{aligned} \tag{20}$$

that is,

$$w^2 + b\, w + \alpha\beta\tau^2 = 0, \qquad b = (\alpha + \beta)\tau - \alpha\beta\, u'v', \tag{21}$$

with roots

$$w_\pm = \tfrac{1}{2}\big(-b \pm \sqrt{b^2 - 4\alpha\beta\tau^2}\big), \qquad w_+ + w_- = -b, \qquad w_+ w_- = \alpha\beta\tau^2.$$

Consider first $u'v' < 0$. Then $-\alpha\beta\, u'v' > 0$, so

$$\begin{aligned}b &> (\alpha + \beta)\tau, \\ b^2 - 4\alpha\beta\tau^2 &> (\alpha + \beta)^2\tau^2 - 4\alpha\beta\tau^2 \\ &= (\alpha - \beta)^2\tau^2 \;\geq\; 0.\end{aligned} \tag{22}$$

Hence both roots are real. Their product $\alpha\beta\tau^2$ is positive and their sum $-b$ is negative, giving $w_- < w_+ < 0$. The four eigenvalues are therefore

$$s = \pm i\sqrt{-w_+}, \qquad s = \pm i\sqrt{-w_-},$$

two imaginary pairs, and their difference exceeds the uncoupled one,

$$w_+ - w_- = \sqrt{b^2 - 4\alpha\beta\tau^2} \;>\; |\alpha - \beta|\,\tau = \big|\alpha\tau - \beta\tau\big|,$$

so the two frequencies are split further apart than $\sqrt{\alpha\tau}$ and $\sqrt{\beta\tau}$. Consider next $u'v' > 0$. Then $b = (\alpha + \beta)\tau - \alpha\beta\, u'v'$ decreases below $(\alpha + \beta)\tau$ as the coupling grows. Once $b^2 < 4\alpha\beta\tau^2$ the discriminant is negative,

$$w_\pm = -\frac{b}{2} \pm \frac{i}{2}\sqrt{4\alpha\beta\tau^2 - b^2},$$

so the roots are complex conjugates off the real axis and the four eigenvalues $s = \pm\sqrt{w_\pm}$ acquire nonzero real parts. For $u'v'$ large enough that $b < 0$ both roots are real and positive (their sum $-b > 0$ and their product $\alpha\beta\tau^2 > 0$). Then $s = +\sqrt{w_+} > 0$ is an eigenvalue with positive real part. The remaining case $b \geq 2\sqrt{\alpha\beta}\,\tau$ requires $\alpha\beta\, u'v' \leq (\sqrt{\alpha} - \sqrt{\beta})^2\tau$, is empty at $\alpha = \beta$, and keeps both roots real and negative with the eigenvalues imaginary. At $\alpha = \beta$ write $g = \sqrt{u'v'}$. Then $b = 2\alpha\tau - \alpha^2 g^2$ and

$$\begin{aligned}b^2 - 4\alpha^2\tau^2 &= \big(2\alpha\tau - \alpha^2 g^2\big)^2 - 4\alpha^2\tau^2 \\ &= \alpha^4 g^4 - 4\alpha^3\tau g^2 \\ &= \alpha^2 g^2\,\big(\alpha^2 g^2 - 4\alpha\tau\big),\end{aligned} \tag{23}$$

so the roots of (21) are perfect squares,

$$\begin{aligned}w_\pm &= \frac{1}{2}\Big(\alpha^2 g^2 - 2\alpha\tau \pm \alpha g\,\sqrt{\alpha^2 g^2 - 4\alpha\tau}\Big) \\ &= \Big(\frac{\alpha g}{2} \pm \frac{1}{2}\sqrt{\alpha^2 g^2 - 4\alpha\tau}\Big)^2.\end{aligned} \tag{24}$$

Taking square roots, the four eigenvalues are

$$s = \pm\Big(\frac{\alpha g}{2} \pm \frac{1}{2}\sqrt{\alpha^2 g^2 - 4\alpha\tau}\Big) = \pm\frac{\alpha g}{2} \pm i\sqrt{\alpha\tau - \alpha^2 g^2/4},$$

where the second form holds for $\alpha g^2 < 4\tau$. Their real parts $\pm\alpha g/2$ are nonzero for every $g > 0$. For $\alpha g^2 \geq 4\tau$ the inner square root is real, so all four eigenvalues are real. □

A.3 Proof of the frequency formula

Section 5.2 tests Lemma 1, proved below for general K. Write $u = Ay$ for player 1's payoff vector under the opponent strategy, $g = u - \lambda c$ for the Lagrangian payoff, and $\chi = x^\top c$ for the spending.

Lemma 4 (Replicator form of the continuous-time limit)**.** *In the continuous-time limit the multiplicative-weights update* (1) *on the payoff vector g becomes the replicator flow*

$$\dot{x}_k \;=\; \eta\, x_k\big(g_k - x^\top g\big), \tag{25}$$

and along it the spending satisfies

$$\dot{\chi} \;=\; \eta \sum_k x_k c_k \big(g_k - x^\top g\big) \;=\; \eta\,\mathrm{Cov}_x(c,g) \;=\; \eta\,\mathrm{Cov}_x(c,u) \;-\; \eta\,\lambda\,\mathrm{Var}_x(c), \tag{26}$$

where Cov_x and Var_x denote covariance and variance under the probability vector x.

Proof. Write the multiplicative-weights update (1) on the payoff vector g as $x_k^+ = x_k e^{\eta g_k}/Z$ with $Z = \sum_j x_j e^{\eta g_j}$. Expanding the exponentials to first order in η,

$$\begin{aligned} Z &= \sum_j x_j\big(1 + \eta g_j + O(\eta^2)\big) = 1 + \eta\, x^\top g + O(\eta^2), \\ x_k^+ &= x_k\big(1 + \eta g_k + O(\eta^2)\big)\big(1 - \eta\, x^\top g + O(\eta^2)\big) \\ &= x_k + \eta\, x_k\big(g_k - x^\top g\big) + O(\eta^2), \end{aligned} \tag{27}$$

so the increment per step is $\eta\, x_k(g_k - x^\top g)$ to first order. Its continuous-time limit is (25). The $O(\eta^2)$ constants depend only on $\max_k |g_k|$. The upper bound on the multiplier keeps $\max_k |g_k|$ at most $\|u\|_\infty + \lambda_{\max}\|c\|_\infty$. The expansion is therefore uniform in k and in the state. Differentiating $\chi = \sum_k x_k c_k$ along (25),

$$\begin{aligned} \dot{\chi} &= \sum_k c_k\, \dot{x}_k \\ &= \eta \sum_k x_k c_k \big(g_k - x^\top g\big) \\ &= \eta\Big(\sum_k x_k c_k g_k - \Big(\sum_k x_k c_k\Big)\Big(\sum_k x_k g_k\Big)\Big) \\ &= \eta\,\big(\mathbb{E}_x[c\,g] - \mathbb{E}_x[c]\,\mathbb{E}_x[g]\big) \\ &= \eta\,\mathrm{Cov}_x(c,g), \end{aligned} \tag{28}$$

which gives the first two equalities of (26). Substituting $g = u - \lambda c$ and using the bilinearity of the covariance,

$$\begin{aligned} \mathrm{Cov}_x(c,g) &= \mathrm{Cov}_x(c, u - \lambda c) \\ &= \mathrm{Cov}_x(c,u) - \lambda\,\mathrm{Cov}_x(c,c) \\ &= \mathrm{Cov}_x(c,u) - \lambda\,\mathrm{Var}_x(c), \end{aligned} \tag{29}$$

which is the last equality of (26). □

Restatement (Lemma 1)**.** *For K actions and cost vector c, hold the payoff-cost covariance and the variance of the cost under the strategy fixed at their values at a fixed point where both budgets are met. Linearizing the multiplier-driven motion of spending there gives $\frac{\mathrm{d}}{\mathrm{d}t}(x^\top c) = -\eta\,\mathrm{Var}_x(c)\,\Delta\lambda$ and $\frac{\mathrm{d}}{\mathrm{d}t}\lambda = \tau\eta\,\Delta(x^\top c)$ (time in steps), a harmonic oscillator whose per-step angular frequency is $\omega = \eta\sqrt{\tau\ \mathrm{Var}_x(c)}$, with $\mathrm{Var}_x(c)$ the variance of the costs under the strategy. At $K = 2$, $c = (1, 0)$, this is $\eta\sqrt{p^*(1-p^*)\,\tau}$.*

Proof. Inside the interval $[0, \lambda_{\max}]$ the multiplier update of (4) is $\lambda^{(t+1)} - \lambda^{(t)} = \tau\eta\,(x_t^\top c - d) = \tau\eta\,(\chi_t - d)$, whose continuous-time limit is

$$\dot{\lambda} = \tau\eta\,(\chi - d).$$

Consider perturbations $(\Delta\chi, \Delta\lambda)$ around the slowly varying fixed point (χ_0, λ_0), holding the payoff-cost covariance $\mathrm{Cov}_x(c,u)$ and the variance $\mathrm{Var}_x(c)$ fixed at their local values. Subtracting (26) at (χ_0, λ_0) from (26) at $(\chi_0 + \Delta\chi, \lambda_0 + \Delta\lambda)$, and likewise for the multiplier equation,

$$\begin{aligned}\Delta\dot{\chi} &= \big[\eta\,\mathrm{Cov}_x(c,u) - \eta\,(\lambda_0 + \Delta\lambda)\;\mathrm{Var}_x(c)\big] - \big[\eta\,\mathrm{Cov}_x(c,u) - \eta\,\lambda_0\;\mathrm{Var}_x(c)\big] \\ &= -\,\eta\;\mathrm{Var}_x(c)\,\Delta\lambda, \qquad (30)\\ \Delta\dot{\lambda} &= \tau\eta\,(\chi_0 + \Delta\chi - d) - \tau\eta\,(\chi_0 - d) \\ &= \tau\eta\,\Delta\chi, \qquad (31)\end{aligned}$$

which are the two linearized equations of the lemma. Differentiating the first once more and substituting the second,

$$\begin{aligned}\Delta\ddot{\chi} &= -\,\eta\;\mathrm{Var}_x(c)\,\Delta\dot{\lambda} \\ &= -\,\eta\;\mathrm{Var}_x(c)\cdot\tau\eta\,\Delta\chi \\ &= -\,\eta^2\tau\;\mathrm{Var}_x(c)\;\Delta\chi, \qquad (32)\end{aligned}$$

the equation $\Delta\ddot{\chi} + \omega^2\,\Delta\chi = 0$ of a harmonic oscillator with

$$\omega^2 = \eta^2\,\tau\;\mathrm{Var}_x(c), \qquad \omega = \eta\sqrt{\tau\;\mathrm{Var}_x(c)},$$

which is (6). Time is in steps, so ω is the per-step frequency of the discrete map to leading order in η. For $K = 2$ with $c = (1,0)$ and $x = (p, 1-p)$,

$$\begin{aligned}\chi &= x^\top c = p, \\ \mathrm{Var}_x(c) &= \mathbb{E}_x[c^2] - \mathbb{E}_x[c]^2 = p - p^2 = p(1-p), \qquad (33)\end{aligned}$$

so at the fixed point $p = p^*$ the frequency is $\omega = \eta\sqrt{p^*(1-p^*)\,\tau}$, the two-action form of Corollary 1. □

Remark 1 (Holding the payoff-cost covariance fixed)**.** *For a fixed opponent the payoff vector $u = Ay$ is constant. At a full-support fixed point all Lagrangian payoff entries equal their mean, $u_k - \lambda^* c_k = g^*$ for all k. Hence $u = g^*\mathbf{1} + \lambda^* c$ and $\mathrm{Cov}_x(c,u) = \lambda^*\,\mathrm{Var}_x(c)$ for every x in the simplex. Holding the payoff-cost covariance fixed is then unnecessary, and (26) reads $\dot{\chi} = -\eta\;\mathrm{Var}_x(c)\,(\lambda - \lambda^*)$ along the whole trajectory. When testing the frequency formula we evaluate $\mathrm{Var}_x(c)$ as a tail average along the trajectory.*

A.4 Energy under the multiplier-first update

Section 4.2 uses Propositions 1 and 2, proved here.

Lemma 5 (The linearized multiplier-first update)**.** *For an isolated oscillator under the multiplier-first update, the linearization at the fixed point in the state $(\delta\ell, \delta\lambda)$ is*

$$L = \begin{pmatrix} 1 - \tau\eta^2\alpha & -\eta \\ \tau\eta\alpha & 1 \end{pmatrix}, \qquad (34)$$

with determinant exactly 1 and trace $2 - \tau\eta^2\alpha$. For $\tau\eta^2\alpha < 4$ its eigenvalues lie exactly on the unit circle with rotation angle $\arccos(1 - \tau\eta^2\alpha/2) = \eta\sqrt{\alpha\tau} + O(\eta^3)$. For the coupled four-dimensional multiplier-first update the determinant is exactly $1 - \eta^2 u' v'\,\sigma'(\ell_1)\sigma'(\ell_2)$.

Proof. Under the multiplier-first update, one step of the isolated oscillator is

$$\begin{aligned}\lambda^+ &= \lambda + \tau\eta\,\big(\sigma(\ell) - p^*\big), \\ \ell^+ &= \ell + \eta\,\big(u - \lambda^+\big), \qquad (35)\end{aligned}$$

the strategy step using the new multiplier. Its fixed point is $\sigma(\ell^*) = p^*$, $\lambda^* = u$. Write $\delta\ell = \ell - \ell^*$, $\delta\lambda = \lambda - \lambda^*$, and expand $\sigma(\ell) = p^* + \alpha\,\delta\ell + O(\delta\ell^2)$ with $\alpha = \sigma'(\ell^*)$:

$$\begin{aligned}\delta\lambda^+ &= \delta\lambda + \tau\eta\alpha\,\delta\ell + O(\delta\ell^2), \\ \delta\ell^+ &= \delta\ell - \eta\,\delta\lambda^+ \\ &= \delta\ell - \eta\,\delta\lambda - \tau\eta^2\alpha\,\delta\ell + O(\delta\ell^2) \\ &= (1 - \tau\eta^2\alpha)\,\delta\ell - \eta\,\delta\lambda + O(\delta\ell^2). \qquad (36)\end{aligned}$$

Collecting the linear terms in the state $(\delta\ell, \delta\lambda)$ gives (34). Its determinant and trace are

$$\begin{aligned}\det L &= (1-\tau\eta^2\alpha)\cdot 1 - (-\eta)\,(\tau\eta\alpha)\\ &= 1-\tau\eta^2\alpha+\tau\eta^2\alpha\\ &= 1, \end{aligned} \tag{37}$$

$$\operatorname{tr} L = (1-\tau\eta^2\alpha)+1 = 2-\tau\eta^2\alpha. \tag{38}$$

A 2×2 matrix with determinant 1 and trace t has characteristic polynomial $\mu^2 - t\mu + 1$, with roots

$$\mu_\pm = \frac{t}{2} \pm \sqrt{\frac{t^2}{4}-1}.$$

For $t \in (-2,2)$ the square root is imaginary, giving $|\mu_\pm|^2 = \frac{t^2}{4} + \left(1-\frac{t^2}{4}\right) = 1$ and $\mu_\pm = e^{\pm i\theta}$ with $\cos\theta = t/2$. Here

$$\frac{t}{2} = 1 - \frac{\tau\eta^2\alpha}{2} \in (-1,1) \quad\Longleftrightarrow\quad 0 < \tau\eta^2\alpha < 4,$$

and the rotation angle is $\theta = \arccos(1-\tau\eta^2\alpha/2)$. Writing $\epsilon = \tau\eta^2\alpha/2$ and using $\cos\theta = 1 - \theta^2/2 + O(\theta^4)$,

$$\begin{aligned} 1-\epsilon &= 1-\frac{\theta^2}{2}+O(\theta^4),\\ \theta^2 &= 2\epsilon + O(\epsilon^2) = \tau\eta^2\alpha + O(\eta^4),\\ \theta &= \eta\sqrt{\alpha\tau}\,\big(1+O(\eta^2)\big) = \eta\sqrt{\alpha\tau} + O(\eta^3). \end{aligned} \tag{39}$$

In the coupled update each multiplier updates first from its own player's strategy and each strategy step uses the opponent's pre-update strategy. With $a_i = \sigma'(\ell_i)$ the linearized step at a general state is

$$\begin{aligned} \delta\lambda_1^+ &= \delta\lambda_1 + \tau\eta\, a_1\,\delta\ell_1, & \delta\ell_1^+ &= \delta\ell_1 - \eta\,\delta\lambda_1^+ + \eta\, u'\, a_2\,\delta\ell_2,\\ \delta\lambda_2^+ &= \delta\lambda_2 + \tau\eta\, a_2\,\delta\ell_2, & \delta\ell_2^+ &= \delta\ell_2 - \eta\,\delta\lambda_2^+ + \eta\, v'\, a_1\,\delta\ell_1, \end{aligned} \tag{40}$$

where the coupling terms come from $u(\sigma(\ell_2))$ and $v(\sigma(\ell_1))$ with slopes u', v'. In the state $(\delta\ell_1, \delta\lambda_1, \delta\ell_2, \delta\lambda_2)$ this map factors as $M = SP$, the multiplier half-step P followed by the strategy half-step S,

$$P = \begin{pmatrix} 1 & 0 & 0 & 0\\ \tau\eta a_1 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & \tau\eta a_2 & 1 \end{pmatrix}, \qquad S = \begin{pmatrix} 1 & -\eta & \eta u' a_2 & 0\\ 0 & 1 & 0 & 0\\ \eta v' a_1 & 0 & 1 & -\eta\\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$M = SP = \begin{pmatrix} 1-\tau\eta^2 a_1 & -\eta & \eta u' a_2 & 0\\ \tau\eta a_1 & 1 & 0 & 0\\ \eta v' a_1 & 0 & 1-\tau\eta^2 a_2 & -\eta\\ 0 & 0 & \tau\eta a_2 & 1 \end{pmatrix}.$$

P is unit lower triangular, so $\det P = 1$. Expanding $\det S$ along its second and fourth rows, which are unit rows,

$$\begin{aligned}\det M &= \det S \cdot \det P\\ &= \det\begin{pmatrix} 1 & \eta u' a_2\\ \eta v' a_1 & 1 \end{pmatrix}\cdot 1\\ &= 1 - \eta^2\, u' v'\, \sigma'(\ell_1)\,\sigma'(\ell_2), \end{aligned} \tag{41}$$

which is the last claim. □

Remark 2 (The two updates compared)**.** *For the same isolated oscillator, the simultaneous update has the linearization of Lemma 6 below with eigenvalue modulus* $(1+\eta^2\alpha\tau)^{1/2} = 1 + \eta^2\alpha\tau/2 + O(\eta^4)$*, against the exact* 1 *of Lemma 5. Both updates therefore share the oscillation frequency to leading order, and with it the frequency formula. Their step-map eigenvalue moduli differ, and that gap is the volume-change difference between the two updates. By the determinant identity (41), all volume change of the coupled multiplier-first update enters through the coupling between the two oscillators.*

Proposition 1 (Multiplier-first update: an isolated oscillator conserves energy)**.** *For an isolated two-action oscillator under the multiplier-first update with $\alpha = p^*(1-p^*)$, the fixed-point linearization has determinant exactly* 1*. For $\tau\eta^2\alpha < 4$ the quadratic $Q = \tau\alpha\,\delta\ell^2 + \tau\eta\alpha\,\delta\ell\,\delta\lambda + \delta\lambda^2$ in the (logit, multiplier) deviation is positive definite and exactly invariant: orbits lie on ellipses and an isolated oscillator never spirals. This holds per isolated oscillator at every K with $p^*(1-p^*)$ replaced by $\mathrm{Var}_x(c)$.*

Proof. Equation (37) is the determinant statement. Let $z = (\delta\ell, \delta\lambda)^\top$. The linearized step is $z^+ = Lz$, and a quadratic form $z^\top Y z$ is exactly invariant if and only if $L^\top Y L = Y$. Write $Y = \begin{pmatrix} y_1 & y_2 \\ y_2 & y_3 \end{pmatrix}$ and abbreviate $a = 1-\tau\eta^2\alpha$, $\nu = \tau\eta\alpha$, so that $L = \begin{pmatrix} a & -\eta \\ \nu & 1 \end{pmatrix}$. Then

$$\begin{aligned} L^\top Y L &= \begin{pmatrix} a & \nu \\ -\eta & 1 \end{pmatrix} \begin{pmatrix} y_1 a + y_2 \nu & -\eta y_1 + y_2 \\ y_2 a + y_3 \nu & -\eta y_2 + y_3 \end{pmatrix} \\ &= \begin{pmatrix} a^2 y_1 + 2a\nu\, y_2 + \nu^2 y_3 & -a\eta\, y_1 + (a - \nu\eta)\, y_2 + \nu\, y_3 \\ -a\eta\, y_1 + (a-\nu\eta)\, y_2 + \nu\, y_3 & \eta^2 y_1 - 2\eta\, y_2 + y_3 \end{pmatrix}, \end{aligned} \tag{42}$$

and $L^\top Y L = Y$ reduces to the three scalar equations

$$\eta^2 y_1 - 2\eta\, y_2 = 0, \tag{43}$$

$$-a\eta\, y_1 + (a - \nu\eta - 1)\, y_2 + \nu\, y_3 = 0, \tag{44}$$

$$(a^2 - 1)\, y_1 + 2a\nu\, y_2 + \nu^2 y_3 = 0. \tag{45}$$

The first of these gives $y_2 = \eta y_1/2$. Substituting into the second, with $a - \nu\eta - 1 = -2\tau\eta^2\alpha$ and $a + \tau\eta^2\alpha = 1$,

$$\begin{aligned} 0 &= -a\eta\, y_1 - 2\tau\eta^2\alpha \cdot \frac{\eta y_1}{2} + \tau\eta\alpha\, y_3 \\ &= -\eta\, y_1 \left(a + \tau\eta^2\alpha\right) + \tau\eta\alpha\, y_3 \\ &= -\eta\, y_1 + \tau\eta\alpha\, y_3, \end{aligned} \tag{46}$$

so $y_3 = y_1/(\tau\alpha)$. The third equation is then satisfied identically, using $a^2 - 1 = -\tau\eta^2\alpha\,(2 - \tau\eta^2\alpha)$:

$$\begin{aligned} (a^2-1)\, y_1 + 2a\nu\, y_2 + \nu^2 y_3 &= -\tau\eta^2\alpha\,(2 - \tau\eta^2\alpha)\, y_1 + a\,\tau\eta^2\alpha\, y_1 + \tau\eta^2\alpha\, y_1 \\ &= \tau\eta^2\alpha\, y_1 \left(-2 + \tau\eta^2\alpha + 1 - \tau\eta^2\alpha + 1\right) \\ &= 0. \end{aligned} \tag{47}$$

Choosing $y_1 = \tau\alpha$ gives the one-dimensional solution family spanned by

$$Y = \begin{pmatrix} \tau\alpha & \tau\eta\alpha/2 \\ \tau\eta\alpha/2 & 1 \end{pmatrix}, \tag{48}$$

an exact identity for every (η, τ, α). The quadratic form

$$\begin{aligned} Q(z) = z^\top Y z &= \tau\alpha\,\delta\ell^2 + 2 \cdot \frac{\tau\eta\alpha}{2}\,\delta\ell\,\delta\lambda + \delta\lambda^2 \\ &= \tau\alpha\,\delta\ell^2 + \tau\eta\alpha\,\delta\ell\,\delta\lambda + \delta\lambda^2 \end{aligned} \tag{49}$$

is the Q of the statement. Its positive definiteness follows from Sylvester's criterion,

$$\begin{aligned} &Y_{11} = \tau\alpha > 0, \\ &\det Y = \tau\alpha \cdot 1 - \left(\frac{\tau\eta\alpha}{2}\right)^2 = \tau\alpha\left(1 - \frac{\tau\eta^2\alpha}{4}\right) > 0 \quad \Longleftrightarrow \quad \tau\eta^2\alpha < 4. \end{aligned} \tag{50}$$

Hence $Q(Lz) = Q(z)$ for every z and Q is positive definite. Orbits of this linearization lie on the ellipses $\{Q = \mathrm{const}\}$ and no radial growth or decay exists at any order. For an isolated oscillator at general K the multiplier uses the spending χ in place of p. By (26) with the payoff-cost covariance held fixed, the linearized strategy equation for χ has coefficient $\mathrm{Var}_x(c)$ in place of $\sigma'(\ell^*) = \alpha$. Everything above applies verbatim with that replacement. □

Remark 3 (Numerical check of Proposition 1)**.** *Numerically, Q is constant to accumulated rounding (10^{-12} relative over 10^6 steps). The full nonlinear multiplier-first update of an isolated oscillator in the interior region shows no secular (steadily accumulating) energy growth over 10^7 steps. The largest deviation of the first integral H over the last 90% of the run does not exceed the largest over the first 10% by more than 5%. By contrast the modulus $(1+\eta^2\alpha\tau)^{1/2} > 1$ per step of the simultaneous update compounds.*

Remark 4 (Energy accounting by region on the two-action multiplier-first update)**.** *We test directly whether the folding argument transfers from the simultaneous update to the multiplier-first update on random two-action games ($\rho = 0$, $\zeta = 0.8$, $\tau = 1$, $\eta = 0.05$, $T = 30{,}000$). Per player the energy is the first integral of the continuous-time limit above its minimum, $E_i = \tau\,\Delta c_i^2[V_i(\ell_i) - V_i(\ell_i^*)] + (\lambda_i \Delta c_i - u_i^*)^2/2$. Δc_i is the cost difference, ℓ_i the logit difference, $V_i(\ell) = \log(1+e^{\ell}) - p_i^*\ell$, and u_i^* the payoff difference at the fixed point. Each step is a step at zero for player i when its projected multiplier is zero and a step away from zero otherwise. We fixed two criteria in advance. Energy growth away from zero transfers if the mean energy change per step away from zero is positive for at least three quarters of the oscillators. The energy drop at a visit to zero holds if the energy after the visit is below the energy before it in at least three quarters of the visits. A first version of this test failed both criteria, at 0.51 and 0.42. It measured the energy against $u_i(q_t)$, the payoff difference at the opponent's current strategy, so energy changes mixed the growth of the oscillation with the motion of that reference. It also filtered out no step at zero, so half of its steps sat at zero, in games whose multiplier never leaves zero. The test reported below evaluates the energy at the fixed u_i^* of the definition above and restricts the population accordingly. Both changes were made after that reading; the two criteria themselves are the ones fixed in advance. The energy drop at a visit to zero holds among the 275 games (of $20{,}000$) whose fixed point exists, whose coupling is competitive, and whose steps at zero are below half of the run for each player. The exit energy is below the entry energy in 0.96 of $4{,}254$ visits (mean log ratio -0.49). Energy growth away from zero does not hold as a population statement. The mean change per step away from zero is positive for 0.59 of the oscillators (median 4×10^{-7}), and 0.49 of the steps away from zero lower the energy. The change over steps away from zero is positive for 0.92 of the oscillators under coordination coupling, whose fixed point is hyperbolic. The strict energy growth of the simultaneous update away from zero does not transfer. This is consistent with Proposition 1.*

Proposition 2 (Multiplier-first update: coupling injects energy)**.** *For the coupled linearized multiplier-first update, the change of the weighted invariant $Q_w = -v'Q_1 + u'Q_2$ vanishes identically at zero coupling. Every energy change is then a coupling cross term. Averaged over independent phases on the invariant ellipses, this change is exactly*

$$\mathbb{E}[\Delta Q_w] = \tfrac{\eta^2}{2}\,\alpha_1\alpha_2\,g_1 g_2\left(\frac{g_2 Q_1}{1-\tau\eta^2\alpha_1/4} + \frac{g_1 Q_2}{1-\tau\eta^2\alpha_2/4}\right) > 0, \qquad g_1 = u' > 0,\ g_2 = -v' > 0, \tag{51}$$

whenever $\tau\eta^2\alpha_i < 4$ for both players, and to leading order in η it is $\frac{\eta^2}{2}\,\alpha_1\alpha_2\,g_1 g_2\,(g_2 Q_1 + g_1 Q_2)$. Every competitive coupling therefore injects energy on average over independent phases on the uncoupled invariant ellipses.

Proof. Write the coupled linearized multiplier-first update of Lemma 5 at the fixed point, where $a_i = \alpha_i$, in the state $z = (\delta\ell_1, \delta\lambda_1, \delta\ell_2, \delta\lambda_2)^\top$ as

$$\begin{aligned} M &= M_0 + \eta\, C, \qquad M_0 = \begin{pmatrix} L_1 & 0 \\ 0 & L_2 \end{pmatrix}, \qquad C = \begin{pmatrix} 0 & C_{12} \\ C_{21} & 0 \end{pmatrix}, \\ C_{12} &= u'\alpha_2\, e_1 e_1^\top, \qquad C_{21} = v'\alpha_1\, e_1 e_1^\top, \end{aligned} \tag{52}$$

where L_i is (34) with $\alpha = \alpha_i$ and $e_1 = (1,0)^\top$. The two coupling entries $\eta u'\alpha_2$ and $\eta v'\alpha_1$ lie in the strategy rows because each strategy update uses the opponent's pre-update strategy. Let Y_i be the isolated-oscillator invariant matrix (48) with $\alpha = \alpha_i$, and set $Y_w = \mathrm{diag}(-v'Y_1,\ u'Y_2)$, with $Q_w = z^\top Y_w z = -v'Q_1 + u'Q_2$. One step changes Q_w by $\Delta Q_w = z^\top \Delta\, z$ with

$$\begin{aligned} \Delta &= M^\top Y_w M - Y_w \\ &= (M_0 + \eta C)^\top Y_w (M_0 + \eta C) - Y_w \\ &= \left(M_0^\top Y_w M_0 - Y_w\right) + \eta\left(C^\top Y_w M_0 + M_0^\top Y_w C\right) + \eta^2\, C^\top Y_w C. \end{aligned} \tag{53}$$

Since M_0 and Y_w are block diagonal and $L_i^\top Y_i L_i = Y_i$ by the proof of Proposition 1,

$$M_0^\top Y_w M_0 - Y_w = \begin{pmatrix} -v'\left(L_1^\top Y_1 L_1 - Y_1\right) & 0 \\ 0 & u'\left(L_2^\top Y_2 L_2 - Y_2\right) \end{pmatrix} = 0, \tag{54}$$

so at $u' = v' = 0$ the increment Δ vanishes identically: each Q_i is conserved and every energy change is a coupling cross term. On an uncoupled invariant ellipse $\{z_i^\top Y_i z_i = Q_i\}$ write $z_i = \sqrt{Q_i}\, Y_i^{-1/2} w_i$ with w_i on the unit circle. Since $L_i^\top Y_i L_i = Y_i$, the matrix $Y_i^{1/2} L_i Y_i^{-1/2}$ is orthogonal. It acts on w_i as the rotation by the angle θ_i of Lemma 5. For irrational $\theta_i/2\pi$ the orbit of w_i is equidistributed on the circle, giving

$$\begin{aligned} \mathbb{E}[w_i w_i^\top] &= \tfrac{1}{2}\, I, \\ \mathbb{E}[z_i z_i^\top] &= Q_i\, Y_i^{-1/2}\, \mathbb{E}[w_i w_i^\top]\, Y_i^{-1/2} = \frac{Q_i}{2}\, Y_i^{-1}. \end{aligned} \tag{55}$$

Treating the two phases as independent makes the cross-block covariance $\mathbb{E}[z_1 z_2^\top]$ zero, so

$$\mathrm{Cov} = \mathbb{E}[z z^\top] = \mathrm{diag}\Big(\frac{Q_1}{2}\, Y_1^{-1},\ \frac{Q_2}{2}\, Y_2^{-1}\Big), \qquad \mathbb{E}[\Delta Q_w] = \mathbb{E}\big[z^\top \Delta\, z\big] = \mathrm{Tr}(\Delta\, \mathrm{Cov}).$$

Insert (53). The first term is zero. In the second, $C^\top Y_w M_0$ and $M_0^\top Y_w C$ are block off-diagonal, because they are the off-diagonal C times block-diagonal factors. A block off-diagonal matrix times the block-diagonal Cov has zero trace, which makes the $O(\eta)$ term vanish exactly. The third term is block diagonal,

$$\begin{aligned} C^\top Y_w C &= \begin{pmatrix} C_{21}^\top\, (u' Y_2)\, C_{21} & 0 \\ 0 & C_{12}^\top\, (-v' Y_1)\, C_{12} \end{pmatrix} \\ &= \begin{pmatrix} u'\,(v'\alpha_1)^2\,(Y_2)_{11}\; e_1 e_1^\top & 0 \\ 0 & -v'\,(u'\alpha_2)^2\,(Y_1)_{11}\; e_1 e_1^\top \end{pmatrix} \\ &= \begin{pmatrix} u' v'^2 \alpha_1^2\, \tau\alpha_2\; e_1 e_1^\top & 0 \\ 0 & -\,v' u'^2 \alpha_2^2\, \tau\alpha_1\; e_1 e_1^\top \end{pmatrix}, \end{aligned} \tag{56}$$

using $(Y_i)_{11} = \tau\alpha_i$ from (48), and the same matrix gives

$$(Y_i^{-1})_{11} = \frac{(Y_i)_{22}}{\det Y_i} = \frac{1}{\tau\alpha_i\,(1 - \tau\eta^2\alpha_i/4)} = \frac{\kappa_i}{\tau\alpha_i}, \qquad \kappa_i = \frac{1}{1 - \tau\eta^2\alpha_i/4} = 1 + O(\eta^2).$$

Therefore

$$\begin{aligned} \mathbb{E}[\Delta Q_w] &= \eta^2\, \mathrm{Tr}\big(C^\top Y_w C\, \mathrm{Cov}\big) \\ &= \eta^2 \Big[u' v'^2 \alpha_1^2\, \tau\alpha_2 \cdot \frac{Q_1}{2}\, (Y_1^{-1})_{11} \;-\; v' u'^2 \alpha_2^2\, \tau\alpha_1 \cdot \frac{Q_2}{2}\, (Y_2^{-1})_{11}\Big] \\ &= \frac{\eta^2}{2}\Big[u' v'^2\, \alpha_1\alpha_2\, \kappa_1 Q_1 - v' u'^2\, \alpha_1\alpha_2\, \kappa_2 Q_2\Big] \\ &= \frac{\eta^2}{2}\, \alpha_1\alpha_2\, u' v'\, \big(v'\kappa_1 Q_1 - u'\kappa_2 Q_2\big). \end{aligned} \tag{57}$$

With $g_1 = u'$ and $g_2 = -v'$, so that $u'v' = -g_1 g_2$ and $v'\kappa_1 Q_1 - u'\kappa_2 Q_2 = -(g_2\kappa_1 Q_1 + g_1\kappa_2 Q_2)$,

$$\frac{\eta^2}{2}\, \alpha_1\alpha_2\, u' v'\, \big(v'\kappa_1 Q_1 - u'\kappa_2 Q_2\big) = \frac{\eta^2}{2}\, \alpha_1\alpha_2\, g_1 g_2\, \big(g_2\kappa_1 Q_1 + g_1\kappa_2 Q_2\big),$$

which is (51), exact under the single model assumption of independent uniform phases. For competitive coupling $g_1, g_2 > 0$ and $\kappa_i > 0$ the expression is positive for all $Q_1, Q_2 > 0$. Here $\kappa_i > 0$ means $\tau\eta^2\alpha_i < 4$, the same condition under which Q_i is positive definite. Setting $\kappa_i = 1$ gives the leading-order form quoted in the statement. A Monte Carlo check over 2×10^5 independent phase pairs reproduces the exact expression within sampling error at $\eta = 0.05$ and at $\eta = 0.6$, $\tau = 10$. There the leading-order form is already 23% low. □

Remark 5 (Resonant refinement of the independent-phase average)**.** *Proposition 2 averages over independent phases. At $\alpha_1 = \alpha_2 = \alpha$ the two uncoupled blocks rotate at the same angle per step, leaving the relative phase φ between them invariant at leading order. Parameterize each*

ellipse as $z_i = \sqrt{Q_i}\, Y^{-1/2} w_i$ *with* w_i *on the unit circle and* w_2 *offset from* w_1 *by* φ*. The cross-block covariance is then* $\mathbb{E}[z_1 z_2^\top] = \frac{1}{2}\sqrt{Q_1 Q_2}\; Y^{-1/2} R(\varphi)^\top Y^{-1/2}$ *with* R *the rotation matrix. Evaluating* $\mathrm{Tr}(\Delta\,\mathrm{Cov})$ *with this block included gives at leading order (the coefficient simplifies exactly),*

$$\mathbb{E}[\Delta Q_w](\varphi) \;=\; \tfrac{\eta^2}{2}\,\alpha^2 g_1 g_2\,(g_2 Q_1 + g_1 Q_2) \;+\; \eta^2 g_1 g_2\,\alpha^{3/2}\sqrt{\tau}\,\sqrt{Q_1 Q_2}\,\sin\varphi. \tag{58}$$

Measured energy growth per step is close to (58) *evaluated at* $\sin\varphi = 1$*, the ceiling of the phase term. Coupled trajectories lock near quadrature (*$\sin\varphi \approx 1$*). Over the* 200 *random linearized competitive systems the independent-phase average runs low by a factor whose first and ninth deciles are* 3.1 *and* 5.0*. The measured energy growth is at the* $\sin\varphi = 1$ *value system by system: measured-to-ceiling deciles* 1.02–1.06 *with median* 1.04 *and rank correlation* 0.94*. Thirty percent of systems exceed the equal-frequency ceiling by more than five percent (top decile of the ratio* 1.06*). Unequal-frequency draws use the geometric-mean* α*. This excess also decreases with the accumulated frequency detuning (rank correlation* -0.27*), consistent with a resonant origin. We verify the phase-dependent formula directly on equal-frequency systems: median relative error* 0.14 *over an interval of eight rotation periods. On the* 300 *trajectories of the full nonlinear two-action multiplier-first update that stay away from zero, the ratios to the independent-phase average are* 2.0–2.2 *(deciles) and are positive in every case.*

Remark 6 (Low-probability actions)**.** *Low-probability actions are actions whose probability has been driven near zero, and regrowth is the growth of their probability.*

A.5 No folding under the simultaneous update

Throughout this subsection, the map is the simultaneous update of Appendix A.1, the Euler step $F = \mathrm{id} + \eta f$ in logit coordinates with f from (7).

Lemma 6 (Linear part of the simultaneous update)**.** *At the fixed point the Jacobian of* F *is* $\left(\begin{smallmatrix} 1 & -\eta \\ \tau\eta\alpha & 1 \end{smallmatrix}\right)$*, with characteristic polynomial* $(1-\mu)^2+\eta^2\alpha\tau$*. Its eigenvalues are exactly* $\mu = 1 \pm i\eta\omega_0$*, of modulus* $(1+\eta^2\omega_0^2)^{1/2} > 1$ *and rotation angle* $\theta(\eta) = \arctan(\eta\omega_0)$*.*

Proof. Differentiating $F = \mathrm{id} + \eta f$ with f from (7),

$$\begin{aligned} DF(\ell,\lambda) &= I + \eta\, Df(\ell,\lambda) \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \eta \begin{pmatrix} 0 & -1 \\ \tau\,\sigma'(\ell) & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -\eta \\ \tau\eta\,\sigma'(\ell) & 1 \end{pmatrix}, \end{aligned} \tag{59}$$

which at the fixed point (ℓ^*, u), where $\sigma'(\ell^*) = \alpha$, is the displayed matrix. Its characteristic polynomial is

$$\begin{aligned} \det\big(\mu I - DF(\ell^*,u)\big) &= \det\begin{pmatrix} \mu-1 & \eta \\ -\tau\eta\alpha & \mu-1 \end{pmatrix} \\ &= (\mu-1)^2 + \tau\eta^2\alpha \\ &= (1-\mu)^2 + \eta^2\alpha\tau, \end{aligned} \tag{60}$$

with roots

$$\mu = 1 \pm i\eta\sqrt{\alpha\tau} = 1 \pm i\eta\omega_0.$$

Hence

$$\begin{aligned} |\mu|^2 &= 1 + \eta^2\omega_0^2 > 1, \\ \tan\theta &= \frac{\mathrm{Im}\,\mu}{\mathrm{Re}\,\mu} = \eta\omega_0, \qquad \theta(\eta) = \arctan(\eta\omega_0), \end{aligned} \tag{61}$$

and in particular every step of the linear part moves the state outward. □

Proposition 3 (Simultaneous update: no folding by saturation or by competitive coupling)**.** *In logit coordinates the continuous-time limit of an isolated oscillator conserves a convex and coercive*

energy H, where coercive means unbounded at infinity. Every Euler step strictly increases H off the fixed point. Hence no periodic orbit and no invariant closed curve exists, and every orbit in the interior region leaves every compact set. For two coupled oscillators and every competitive coupling $u'v' < 0$, a weighted first integral H_w has the same properties at any coupling strength. For coordination coupling $u'v' > 0$ the fixed point is hyperbolic, and no conserved energy with a nondegenerate minimum exists.

Proof. Define

$$H(\ell, \lambda) = \tau\Big[\log(1 + e^{\ell}) - \frac{\zeta}{2}\,\ell\Big] + \frac{1}{2}\,(\lambda - u)^2.$$

Using $\frac{d}{d\ell}\log(1 + e^{\ell}) = \sigma(\ell)$, its gradient and Hessian are

$$\begin{aligned} \nabla H(\ell, \lambda) &= \Big(\tau\big(\sigma(\ell) - \tfrac{\zeta}{2}\big),\ \lambda - u\Big), \\ \nabla^2 H(\ell, \lambda) &= \operatorname{diag}\big(\tau\,\sigma'(\ell),\ 1\big) \succ 0, \end{aligned} \tag{62}$$

and along the field f of (7)

$$\begin{aligned} \nabla H \cdot f &= \tau\big(\sigma(\ell) - \tfrac{\zeta}{2}\big)\,(u - \lambda) + (\lambda - u)\,\tau\big(\sigma(\ell) - \tfrac{\zeta}{2}\big) \\ &= 0, \end{aligned} \tag{63}$$

so H is a first integral (a conserved quantity) of the continuous-time limit $\dot{x} = f(x)$. The map $F = \mathrm{id} + \eta f$ is its explicit Euler step. Taylor's formula with integral remainder, applied to $t \mapsto H(x + t\eta f)$ on $[0, 1]$, gives for any state $x = (\ell, \lambda)$

$$\begin{aligned} H(x + \eta f) - H(x) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\, H(x + t\eta f)\, dt \\ &= \eta\, \nabla H(x) \cdot f + \int_0^1 (1 - t)\, \frac{d^2}{dt^2}\, H(x + t\eta f)\, dt \\ &= \eta\, \nabla H \cdot f + \eta^2 \int_0^1 (1 - t)\, f^{\top} \nabla^2 H\big(x + t\eta f\big)\, f\, dt, \end{aligned} \tag{64}$$

and with $\nabla H \cdot f = 0$ the first term vanishes:

$$H(Fx) - H(x) = \eta^2 \int_0^1 (1 - t)\, \Big[\tau\, \sigma'\big(\ell + t\eta f_1\big)\, f_1^2 + f_2^2\Big]\, dt \ \geq\ 0.$$

Every Euler step satisfies $H \circ F \geq H$. If equality holds, the nonnegative continuous integrand vanishes for all $t \in [0, 1]$. The λ-component gives

$$f_2 = \tau\big(\sigma(\ell) - \tfrac{\zeta}{2}\big) = 0 \implies \sigma(\ell) = \tfrac{\zeta}{2} = p^* \implies \ell = \ell^*,$$

and since $\sigma' > 0$ everywhere the ℓ-component then gives

$$f_1 = u - \lambda = 0 \implies \lambda = u.$$

Equality holds only at the fixed point (ℓ^*, u), and every step off the fixed point strictly increases H. H is strictly convex by the Hessian above, since $\sigma' > 0$. It is also coercive: with $0 < \zeta/2 < 1$ (Assumption 1),

$$\begin{aligned} \log(1 + e^{\ell}) - \frac{\zeta}{2}\,\ell &= \Big(1 - \frac{\zeta}{2}\Big)\,\ell + \log(1 + e^{-\ell}) \ \longrightarrow\ +\infty \qquad (\ell \to +\infty), \\ \log(1 + e^{\ell}) - \frac{\zeta}{2}\,\ell &= -\frac{\zeta}{2}\,\ell + \log(1 + e^{\ell}) \ \longrightarrow\ +\infty \qquad (\ell \to -\infty), \end{aligned} \tag{65}$$

and $(\lambda - u)^2/2 \to \infty$ as $|\lambda| \to \infty$. The fixed point is the unique critical point of H and therefore its global minimum. No periodic orbit can exist: if $F^n x = x$ for some $n \geq 1$ and x not the fixed point, then

$$H(x) < H(Fx) < H(F^2 x) < \cdots < H(F^n x) = H(x),$$

a contradiction. No invariant closed curve Γ can exist either: H attains its maximum on the compact set Γ at some x_0, and invariance gives $F(x_0) \in \Gamma$, so

$$H(x_0) \ \geq\ H(F(x_0)) \ >\ H(x_0)$$

unless x_0 is the fixed point. That is the global minimum of H and so can be the maximizer only if Γ is the fixed point itself. Finally, $H_n = H(F^n x)$ is strictly increasing along every non-fixed orbit. If it were bounded it would converge to some H_∞, and the orbit would stay in the compact sublevel set $\{H \le H_\infty\}$. Any accumulation point $x_\infty = \lim_k F^{n_k} x$ would then satisfy, by continuity of F and H,

$$H(Fx_\infty) - H(x_\infty) = \lim_k \big(H_{n_k+1} - H_{n_k}\big) = 0,$$

that is, x_∞ would be the fixed point, whose H-value is $\min H \le H_0 < H_1 \le H_\infty = H(x_\infty)$: a contradiction. Hence $H_n \to \infty$ and every trajectory in the interior region leaves every compact set, which proves the isolated-oscillator claims.

With both budgets met with equality, write $(p_1, p_2) = (p, q)$, $\delta p_i = p_i - p^*$ and $r_i = \lambda_i - \lambda_i^*$, with λ_i^* the fixed-point multipliers of Lemma 2. Since u is affine with slope u', $u(p_2) = u(p^*) + u'\delta p_2 = \lambda_1^* + u'\delta p_2$ and likewise $v(p_1) = \lambda_2^* + v'\delta p_1$. The coupled field in the state $(\ell_1, r_1, \ell_2, r_2)$ is then

$$\begin{aligned} f &= \big(u(p_2) - \lambda_1,\ \tau\,(p_1 - p^*),\ v(p_1) - \lambda_2,\ \tau\,(p_2 - p^*)\big) \\ &= \big(u'\delta p_2 - r_1,\ \tau\delta p_1,\ v'\delta p_1 - r_2,\ \tau\delta p_2\big). \end{aligned} \tag{66}$$

The map is the Euler step of f, and $H_i = \tau\big[\log(1+e^{\ell_i}) - \frac{\zeta}{2}\,\ell_i\big] + \frac{1}{2}r_i^2$ is the first integral of oscillator i with the payoff difference held at its value at the fixed point. With $\nabla H_1 = (\tau\delta p_1, r_1, 0, 0)$ and $\nabla H_2 = (0, 0, \tau\delta p_2, r_2)$, since $\tau\,(\sigma(\ell_i) - \zeta/2) = \tau\delta p_i$,

$$\begin{aligned} \nabla H_1 \cdot f &= \tau\delta p_1\,(u'\delta p_2 - r_1) + r_1\,\tau\delta p_1 = \tau u'\,\delta p_1\delta p_2, \\ \nabla H_2 \cdot f &= \tau\delta p_2\,(v'\delta p_1 - r_2) + r_2\,\tau\delta p_2 = \tau v'\,\delta p_1\delta p_2\,: \end{aligned} \tag{67}$$

the self-terms cancel exactly as in the isolated oscillator and only the coupling contributes. Hence $H_w = -v'H_1 + u'H_2$ satisfies

$$\nabla H_w \cdot f = -v'\,\tau u'\,\delta p_1\delta p_2 + u'\,\tau v'\,\delta p_1\delta p_2 = 0 \tag{68}$$

for every coupling: H_w is conserved by the coupled continuous-time limit. For $u'v' < 0$ the weights $-v'$ and u' share the sign of u', so

$$\operatorname{sign}(u')\,H_w = |v'|\,H_1 + |u'|\,H_2$$

is a positive combination of the convex, coercive H_i (equivalently, relabel so that $u' > 0 > v'$). It is strictly convex and coercive with a unique critical point at the fixed point. Identity (64) applies verbatim with

$$\nabla^2\big[\operatorname{sign}(u')\,H_w\big] = \operatorname{diag}\big(|v'|\tau\sigma'(\ell_1),\ |v'|,\ |u'|\tau\sigma'(\ell_2),\ |u'|\big) \succ 0,$$

so every Euler step increases $\operatorname{sign}(u')H_w$, and equality forces the integrand to vanish:

$$\tau\delta p_i = 0 \implies \ell_i = \ell^*, \qquad \text{and then} \qquad u'\delta p_2 - r_1 = 0,\ v'\delta p_1 - r_2 = 0 \implies r_1 = r_2 = 0,$$

the fixed point. All three conclusions above (no periodic orbit, no invariant closed curve, unbounded growth) follow word for word with H replaced by $\operatorname{sign}(u')H_w$. This proves the competitive-coupling claim.

For $u'v' > 0$ both players share $p^* = \zeta/2$ in this continuous-time limit, giving $\alpha = \beta$. Corollary 1 then gives the eigenvalues of the linearized continuous-time limit in closed form: with $g = \sqrt{u'v'}$ and $\Omega = \sqrt{\alpha\tau - \alpha^2 g^2/4}$,

$$s \in \Big\{ \pm\frac{\alpha g}{2} \pm i\,\Omega \Big\} \quad (\alpha g^2 < 4\tau), \qquad s \in \Big\{ \pm\frac{\alpha g}{2} \pm \frac{1}{2}\sqrt{\alpha^2 g^2 - 4\alpha\tau} \Big\} \subset \mathbb{R}\setminus\{0\} \quad (\alpha g^2 \ge 4\tau);$$

every eigenvalue has real part $\pm\alpha g/2 \neq 0$ or is real and nonzero, which makes the fixed point hyperbolic for every coupling strength $g > 0$. Suppose a C^2 function G were conserved by the continuous-time limit near the fixed point, $\nabla G(x) \cdot f(x) \equiv 0$, with a critical point there. Then $\nabla G(x) = \Sigma x + o(|x|)$ with Σ the Hessian of G at the fixed point, while $f(x) = Jx + O(|x|^2)$. Expanding at quadratic order,

$$\begin{aligned} 0 \equiv \nabla G(x) \cdot f(x) &= x^\top \Sigma J\, x + o(|x|^2) \\ \implies \quad x^\top\big(\Sigma J + J^\top\Sigma\big)\,x = 0 \quad &\text{for all } x \quad \implies \quad J^\top\Sigma + \Sigma J = 0, \end{aligned} \tag{69}$$

the Lyapunov equation for Σ, since a symmetric matrix whose quadratic form vanishes identically is zero. For an eigenvector $Jw = sw$, multiply the Lyapunov equation by $\bar{w}^\top$ on the left and by w on the right, using that J is real:

$$\begin{aligned} 0 &= \bar{w}^\top J^\top \Sigma\, w + \bar{w}^\top \Sigma J\, w \\ &= (J\bar{w})^\top \Sigma\, w + s\, \bar{w}^\top \Sigma\, w \\ &= \bar{s}\, \bar{w}^\top \Sigma\, w + s\, \bar{w}^\top \Sigma\, w \\ &= (s + \bar{s})\, \bar{w}^\top \Sigma\, w, \end{aligned} \tag{70}$$

and $s + \bar{s} = 2\,\mathrm{Re}\, s \neq 0$ by hyperbolicity forces $\bar{w}^\top \Sigma w = 0$. A nondegenerate minimum of G at the fixed point means $\Sigma \succ 0$, and writing the nonzero eigenvector as $w = w_r + i\, w_i$ with real w_r, w_i,

$$\begin{aligned} \bar{w}^\top \Sigma\, w &= (w_r - i\, w_i)^\top \Sigma\, (w_r + i\, w_i) \\ &= w_r^\top \Sigma\, w_r + w_i^\top \Sigma\, w_i + i\, \big(w_r^\top \Sigma\, w_i - w_i^\top \Sigma\, w_r\big) \\ &= w_r^\top \Sigma\, w_r + w_i^\top \Sigma\, w_i \;>\; 0, \end{aligned} \tag{71}$$

since Σ is symmetric and at least one of w_r, w_i is nonzero. The two displays contradict each other, so for every $g > 0$ no conserved function has a nondegenerate minimum at the fixed point. This proves the coordination-coupling claim. For the symmetric coupling $u' = v' = g$ the in-phase diagonal $\{\ell_1 = \ell_2,\ r_1 = r_2\}$ is exactly invariant, because the field above takes equal values in the two oscillators there. On it, with $\delta p_1 = \delta p_2 = \delta p$, the energy $\tilde{H} = H_1$ satisfies

$$\frac{d\tilde{H}}{dt} = \nabla H_1 \cdot f = \tau u'\, \delta p_1 \delta p_2 = \tau g\, \delta p^2 \;\geq\; 0,$$

so even the continuous-time limit expands monotonically along this invariant diagonal. □

A.6 The criterion for reaching zero

Section 4.2 uses Lemma 7 below.

Lemma 7 (Reaching zero)**.** *Let* $\mathrm{viol}_s = x_s^\top c_x - d_x$ *be player* 1*'s violation at step* s*, with* x_s *the strategy vector at that step, and let* $\lambda_t > 0$*. If the running average violation over the stretch of* W *steps starting at* t *satisfies* $\frac{1}{W}\sum_{s=t}^{t+W-1} \mathrm{viol}_s < -\lambda_t/(\tau\eta W)$*, then the multiplier reaches the bound* $\lambda = 0$ *at some step of that stretch.*

Proof. Here the bound $\lambda = 0$ is the lower bound of the interval $[0, \lambda_{\max}]$ and the projection maps onto that interval. Write the unprojected update as $\tilde{\lambda}_{s+1} = \lambda_s + \tau\eta\, \mathrm{viol}_s$, so that

$$\lambda_{s+1} = \min\big(\max(\tilde{\lambda}_{s+1}, 0),\ \lambda_{\max}\big) \;\leq\; \max(\tilde{\lambda}_{s+1}, 0).$$

If zero is not hit at step s, then $\tilde{\lambda}_{s+1} > 0$ and

$$\lambda_{s+1} \;\leq\; \tilde{\lambda}_{s+1} \;=\; \lambda_s + \tau\eta\, \mathrm{viol}_s,$$

because the upper bound only lowers the value further. Suppose zero is never reached during the stretch $s = t, \ldots, t + W - 1$. Applying the inequality at every step of the stretch and chaining,

$$\begin{aligned} \lambda_{t+W} &\leq \lambda_{t+W-1} + \tau\eta\, \mathrm{viol}_{t+W-1} \\ &\leq \lambda_{t+W-2} + \tau\eta\, \big(\mathrm{viol}_{t+W-2} + \mathrm{viol}_{t+W-1}\big) \\ &\;\;\vdots \\ &\leq \lambda_t + \tau\eta \sum_{s=t}^{t+W-1} \mathrm{viol}_s. \end{aligned} \tag{72}$$

The hypothesis gives

$$\frac{1}{W}\sum_{s=t}^{t+W-1} \mathrm{viol}_s < -\frac{\lambda_t}{\tau\eta W} \quad\Longrightarrow\quad \tau\eta \sum_{s=t}^{t+W-1} \mathrm{viol}_s < -\lambda_t \quad\Longrightarrow\quad \lambda_{t+W} < \lambda_t - \lambda_t = 0,$$

contradicting $\lambda_{t+W} \geq 0$. Hence zero is reached within the stretch. □

The supplement checks the implication of Lemma 7 on simulated trajectories, and Table A35 records how often trajectories reach zero.

A.7 THE VISIT TO ZERO AND THE PERIODIC ORBIT

The periodic orbit that Section 4.2 quotes has a closed form under Assumption 3. A visit to zero is a stretch of steps during which the multiplier stays at zero. The periodic orbit studied here visits zero once per period. The kink is the exact point $\lambda = 0$ of the projection. Throughout, the oscillator is in logit coordinates (ℓ, λ) with $p = \sigma(\ell)$, $p^* = \zeta/2$, and $\ell^* = \sigma^{-1}(p^*)$. The first integral of Proposition 3 is written as $H = \tau V(\ell) + (\lambda - u)^2/2$ with potential $V(\ell) = \log(1 + e^\ell) - p^*\ell$, whose minimum is at ℓ^*. We write $E = H - \tau V(\ell^*)$ for the energy above the minimum of H, $\Delta V(\ell) = V(\ell) - V(\ell^*) \geq 0$ for the potential above its minimum, and $\omega_0 = \sqrt{p^*(1-p^*)\,\tau}$ for the frequency of the linearized continuous-time limit.

Assumption 3 (Periodic-orbit setting)**.** *Here the payoff difference is a constant $u > 0$. The map is the simultaneous update of Appendix A.1 with the projection of (4) applied to the multiplier. The upper bound satisfies $\lambda_{\max} > 2u$ together with the quantitative condition $u^2/2 + \tau(\eta u)^2/8 + \Delta E_{\text{loop}} < (\lambda_{\max} - u)^2/2$, where ΔE_{loop} is the one-revolution energy gain of (80) below. At the ensemble values $\lambda_{\max} = 50$, $\eta = 0.05$, $\tau \leq 10$ the left side stays below 20 for $u \leq 5$ while the right side exceeds 1000.*

The statements below about the side on which the orbit reaches zero, the visit length, and the energy drop at a visit to zero are exact for the discrete map.

Proposition 4 (Entry, visit to zero, and energy drop)**.** *Under Assumption 3: (i) the projection activates only on the side $\ell < \ell^*$ where the multiplier decreases, from some $\lambda \in [0, \tau\eta p^*)$, as soon as $E > u^2/2 + \tau(\eta u)^2/8$ on that side; (ii) with ℓ_{in} the logit of the first state with $\lambda = 0$, the visit to zero has length exactly*

$$n(\ell_{\text{in}}) = \left\lceil \frac{\ell^* - \ell_{\text{in}}}{\eta u} \right\rceil, \tag{73}$$

and ends at $(\ell_{\text{out}}, 0)$ with $\ell_{\text{out}} \in [\ell^*, \ell^* + \eta u)$; (iii) with E_{in} the energy at entry, the energy change across the visit to zero is

$$\Delta H_{\text{sojourn}} = \tau\big[V(\ell_{\text{out}}) - V(\ell_{\text{in}})\big] = -\Big(E_{\text{in}} - \frac{u^2}{2}\Big) + \tau\Delta V(\ell_{\text{out}}), \qquad 0 \leq \tau\Delta V(\ell_{\text{out}}) \leq \frac{\tau(\eta u)^2}{8}, \tag{74}$$

so the exit energy satisfies $u^2/2 \leq E_{\text{out}} \leq u^2/2 + \tau(\eta u)^2/8$ independently of E_{in}; (iv) the return map in entry energy composes one revolution of the flow away from zero with the energy drop (74); in one return it maps all of $\{E > u^2/2\}$ below the threshold $(\lambda_{\max} - u)^2/2$ for reaching the upper bound into an interval of relative width at most $\tau\eta^2/4$.

Proof. The unprojected multiplier step of the simultaneous update from a state (ℓ, λ) with $\lambda \geq 0$ is $\tilde{\lambda}^+ = \lambda + \tau\eta\,(\sigma(\ell) - p^*)$, and the projection at zero activates precisely when $\tilde{\lambda}^+ < 0$:

$$\lambda + \tau\eta\,\big(\sigma(\ell) - p^*\big) < 0 \implies \tau\eta\,\big(\sigma(\ell) - p^*\big) < -\lambda \leq 0 \implies \sigma(\ell) < p^* \implies \ell < \ell^*,$$

so entry always happens on the side where the multiplier decreases. An interior multiplier step moves λ down by at most

$$\tau\eta\,\big(p^* - \sigma(\ell)\big) < \tau\eta\,p^*,$$

so a state with $\lambda \geq \tau\eta p^*$ cannot reach zero in one step: the orbit cannot skip the strip $0 \leq \lambda < \tau\eta p^*$, and the projection acts from inside it. On $\lambda = 0$ the multiplier part of H is $u^2/2$, so the conservative level set $\{E = e\}$ meets that line at the logits ℓ with

$$e = \frac{u^2}{2} + \tau\,\Delta V(\ell).$$

This has a solution $\ell < \ell^*$ exactly when $e > u^2/2$, because ΔV decreases from $+\infty$ to 0 on $(-\infty, \ell^*]$. The outward spiral of Proposition 3 raises E every step. Once $E > u^2/2 + \tau(\eta u)^2/8$ the projection does activate, which we show by contradiction. Suppose it never does from some state with $\ell < \ell^*$ and $\lambda \geq 0$ onwards, so that every later step is unprojected and E rises at every one of them. While $\ell < \ell^*$ the multiplier falls by $\tau\eta\,(p^* - \sigma(\ell)) > 0$ per step. While $\lambda \geq u$ the logit does not increase, so that decrease is bounded below. Then $\lambda < u$ after finitely many steps, with ℓ still below ℓ^*. From there the logit rises by $\eta(u - \lambda) > 0$ per step while the multiplier keeps falling, so ℓ first reaches ℓ^* at some step, entering $[\ell^*, \ell^* + \eta u)$ from a state with $\lambda < u$. At

that step $E = \tau\,\Delta V(\ell) + (u-\lambda)^2/2$ exactly, and $\tau\,\Delta V(\ell) \le \tau(\ell-\ell^*)^2/8 < \tau(\eta u)^2/8$ because $V'' = \sigma' \le 1/4$. Hence $(u-\lambda)^2/2 > u^2/2$ and $\lambda < 0$, contradicting $\lambda \ge 0$. The projection step replaces λ by 0 and changes the multiplier part of H by

$$\begin{aligned}\frac{1}{2}\,(0-u)^2 - \frac{1}{2}\,(\lambda-u)^2 &= \frac{u^2}{2} - \frac{\lambda^2}{2} + u\lambda - \frac{u^2}{2} \\ &= u\lambda - \frac{\lambda^2}{2} = \lambda\left(u - \frac{\lambda}{2}\right) \;\ge\; 0 \qquad (\lambda \le 2u),\end{aligned} \tag{75}$$

a bounded entry correction below $\tau\eta p^* u$ that the subsequent visit to zero discards along with the rest of the potential excess. This proves (i). While $\lambda = 0$ and $\sigma(\ell) < p^*$, the step of the simultaneous update is affine:

$$\begin{aligned}\lambda_{k+1} &= \max\big(0 + \tau\eta\,(\sigma(\ell_k) - p^*),\; 0\big) = 0, \\ \ell_{k+1} &= \ell_k + \eta\,(u - \lambda_k) = \ell_k + \eta u,\end{aligned} \tag{76}$$

so by induction from the entry state $(\ell_{\mathrm{in}}, 0)$,

$$\ell_k = \ell_{\mathrm{in}} + k\,\eta u, \qquad \lambda_k = 0 \qquad (k = 0, 1, \dots, n).$$

The projection stays active while $\sigma(\ell_k) < p^*$, that is while $\ell_k < \ell^*$, and the visit to zero ends at the first k with

$$\ell_{\mathrm{in}} + k\,\eta u \;\ge\; \ell^* \quad\Longleftrightarrow\quad k \;\ge\; \frac{\ell^* - \ell_{\mathrm{in}}}{\eta u}, \qquad \text{that is} \qquad n(\ell_{\mathrm{in}}) = \left\lceil \frac{\ell^* - \ell_{\mathrm{in}}}{\eta u} \right\rceil,$$

which is (73). The exit state is $(\ell_{\mathrm{out}}, 0)$ with

$$\ell_{\mathrm{out}} = \ell_{\mathrm{in}} + n\,\eta u \in [\ell^*,\; \ell^* + \eta u),$$

since n is the least integer with $\ell_n \ge \ell^*$ while $\ell_{n-1} < \ell^*$. After it the multiplier leaves zero as soon as $\sigma(\ell) > p^*$, at the next step when $\ell_{\mathrm{out}} > \ell^*$ and one step later when $\ell_{\mathrm{out}} = \ell^*$. This proves (ii). During the visit to zero the multiplier part of H is frozen at $\frac{1}{2}(0-u)^2 = u^2/2$. Per step at zero, with $V'(\ell) = \sigma(\ell) - p^*$,

$$\begin{aligned}\Delta H &= \tau\big[V(\ell + \eta u) - V(\ell)\big] \\ &= \tau \int_{\ell}^{\ell+\eta u} \big(\sigma(t) - p^*\big)\,dt \\ &= \tau\eta u\big(\sigma(\ell) - p^*\big) + \frac{\tau\eta^2 u^2}{2}\,\sigma'(\ell) + O(\eta^3).\end{aligned} \tag{77}$$

On every step at zero except possibly the exit step, $\ell + \eta u \le \ell^*$. The integrand is then negative on the whole interval and $\Delta H < 0$: the steps at zero dissipate while the interior region adds energy. Summing (77) over the visit to zero telescopes exactly,

$$\begin{aligned}\Delta H_{\mathrm{sojourn}} &= \sum_{k=0}^{n-1} \tau\big[V(\ell_{k+1}) - V(\ell_k)\big] = \tau\big[V(\ell_{\mathrm{out}}) - V(\ell_{\mathrm{in}})\big] \\ &= \tau\big[V(\ell_{\mathrm{out}}) - V(\ell^*)\big] - \tau\big[V(\ell_{\mathrm{in}}) - V(\ell^*)\big] \\ &= \tau\,\Delta V(\ell_{\mathrm{out}}) - \Big(E_{\mathrm{in}} - \frac{u^2}{2}\Big),\end{aligned} \tag{78}$$

where the last line uses $E_{\mathrm{in}} = H(\ell_{\mathrm{in}}, 0) - \tau V(\ell^*) = \tau\big[V(\ell_{\mathrm{in}}) - V(\ell^*)\big] + u^2/2$. This is (74). A bound on the remainder follows from $\sigma' \le 1/4$ and $0 \le \ell_{\mathrm{out}} - \ell^* < \eta u$:

$$\begin{aligned}0 \;\le\; \tau\,\Delta V(\ell_{\mathrm{out}}) &= \tau \int_{\ell^*}^{\ell_{\mathrm{out}}} \big(\sigma(t) - \sigma(\ell^*)\big)\,dt \\ &\le \tau \int_{\ell^*}^{\ell_{\mathrm{out}}} \frac{t - \ell^*}{4}\,dt \\ &= \frac{\tau\,(\ell_{\mathrm{out}} - \ell^*)^2}{8} \;<\; \frac{\tau\,(\eta u)^2}{8}.\end{aligned} \tag{79}$$

The exit energy is therefore

$$E_{\text{out}} = H(\ell_{\text{out}}, 0) - \tau V(\ell^*) = \frac{u^2}{2} + \tau\,\Delta V(\ell_{\text{out}}) \in \Big[\frac{u^2}{2},\ \frac{u^2}{2} + \frac{\tau(\eta u)^2}{8}\Big],$$

independent of how large E_{in} was: the visit to zero removes the entire accumulated potential excess, which proves (iii). The upper bound $\lambda = \lambda_{\max}$ admits the mirror argument with the multiplier part frozen at $\frac{1}{2}(\lambda_{\max} - u)^2$. For $\lambda_{\max} > 2u$ (Assumption 3),

$$\frac{1}{2}\,(\lambda_{\max} - u)^2 > \frac{1}{2}\,u^2,$$

so the lower bound is reached at smaller energy. After one visit to zero the energy before the next visit is at most the exit energy $u^2/2 + \tau(\eta u)^2/8$ plus the one-revolution gain ΔE_{loop} of (80). The quantitative condition in Assumption 3 keeps that sum below the threshold for reaching the upper bound, so the orbit never again reaches the upper bound. The condition $\lambda_{\max} > 2u$ alone would not suffice when the upper bound sits just above $2u$. From an exit state the orbit spirals once around (ℓ^*, u), to leading order along the conservative level set $\{E = u^2/2\}$. By (64) with $\nabla H \cdot f = 0$, one step away from zero gains

$$\Delta H = \frac{\eta^2}{2}\, f^\top \nabla^2 H\, f + O(\eta^3) = \frac{\eta^2}{2}\Big[\tau\,\sigma'(\ell)\,(u-\lambda)^2 + \tau^2\,\big(\sigma(\ell) - p^*\big)^2\Big] + O(\eta^3),$$

and summing over the steps of one revolution of the flow, with $\eta \sum_{\text{steps}}(\cdot) \to \oint (\cdot)\, dt$,

$$\Delta E_{\text{loop}} = \frac{\eta}{2} P + O(\eta^2), \qquad P = \oint_{E=u^2/2} \big[\tau\sigma'(\ell)(u-\lambda)^2 + \tau^2(\sigma(\ell) - p^*)^2\big]\, dt. \tag{80}$$

In the harmonic limit $V(\ell) \approx V(\ell^*) + \frac{\alpha}{2}\,\delta\ell^2$, the level set $\{E = u^2/2\}$ is the ellipse $\delta\ell = \frac{u}{\omega_0}\cos\phi$, $\delta\lambda = u\sin\phi$ traversed at $|\dot\phi| = \omega_0$. On it, with $\tau\alpha = \omega_0^2$,

$$\begin{aligned}
\tau\,\sigma'(\ell)\,(u-\lambda)^2 + \tau^2\,\big(\sigma(\ell) - p^*\big)^2 &\approx \tau\alpha\,\delta\lambda^2 + \tau^2\alpha^2\,\delta\ell^2 \\
&= \omega_0^2 u^2 \sin^2\phi + \omega_0^4\,\frac{u^2}{\omega_0^2}\cos^2\phi = \omega_0^2 u^2, \\
P &= \int_0^{2\pi/\omega_0} \omega_0^2 u^2\, dt = 2\pi\,\omega_0 u^2, \\
\Delta E_{\text{loop}} &= \frac{\eta}{2}\cdot 2\pi\,\omega_0 u^2 = \frac{u^2}{2}\cdot 2\pi\eta\omega_0,
\end{aligned} \tag{81}$$

matching the per-step energy growth of Remark 2: the modulus $(1 + \eta^2\omega_0^2)^{1/2}$ per step multiplies E by $1 + \eta^2\omega_0^2$, and over the $2\pi/(\eta\omega_0)$ steps of a revolution this compounds to the relative gain $2\pi\eta\omega_0$. E grows past the threshold of (i) during the revolution, so entry is guaranteed within it. At most one visit can occur in a revolution: a visit ends at $\ell_{\text{out}} \geq \ell^*$ by (ii), the projection acts only at $\ell < \ell^*$ by (i), and the logit moves left only from a state with $\lambda > u$. Therefore the orbit must pass around (ℓ^*, u) before it can enter again. Hence the attractor has exactly one visit to zero per revolution. After the energy growth and the energy drop, the return map in entry energy sends every $E_{\text{in}} > u^2/2$ below the threshold for reaching the upper bound to

$$E_{\text{out}} \in \Big[\frac{u^2}{2},\ \frac{u^2}{2} + \frac{\tau(\eta u)^2}{8}\Big), \qquad \frac{\tau(\eta u)^2/8}{u^2/2} = \frac{\tau\eta^2}{4},$$

an interval of relative width at most $\tau\eta^2/4$ in a single return, which proves (iv). The residual dynamics lives only in the entry angle, which advances by a near-rigid rotation (eigenvalues $1 \pm i\eta\omega_0$ of modulus $1 + O(\eta^2)$ by Lemma 6). □

Remark 7 (Stability of the periodic orbit)**.** *Proposition 4 gives the energy drop at a visit to zero. The entry angle advances by a near-rigid rotation with an $O(\eta)$ correction from the visit to zero, so to leading order the return map has derivative $1 + O(\eta) > 0$. Two further properties give the isolated oscillator a non-positive top exponent: an attracting invariant circle with an orientation-preserving C^1 return map, and a generic periodic set at rational rotation number. We assume that the contracted annulus closes into an attracting invariant circle on which the return map (the map from one*

visit to zero to the next) restricts to an orientation-preserving C^1 circle diffeomorphism. We also assume that in the rational-rotation-number case the periodic set is generic: finitely many alternating attracting and repelling orbits rather than intervals of periodic points. Under both properties the rational-versus-irrational dichotomy of the rotation number applies (Katok & Hasselblatt, 1995). Lebesgue-almost-every entry angle converges to an attracting periodic orbit of the return map with along-circle exponent ≤ 0. At irrational rotation number a C^1 circle diffeomorphism has a single invariant probability measure. This unique ergodicity forces the along-circle exponent to zero. With the transverse contraction, the top exponent along Lebesgue-almost-every trajectory is non-positive. The scan over 54 parameter settings finds no exponent above the resolution interval and matches the predictions for the periodic orbit to within 10%. Positive finite-time exponents need two coupled oscillators that each visit zero. As coupling grows the two-frequency torus breaks sharply through overlapping mode-locking regions (top exponent jumping three orders of magnitude within one scan step). An ensemble intervention distinguishes the two roles of the bound at zero: the softplus reparameterization retains the peak of the separating fraction while emptying the class of games on the periodic orbit. Folding needs the low-multiplier region, and only the periodic orbits need the kink itself (Appendix D.5).

Remark 8 (Predictions of the periodic orbit and their measurement)**.** *To leading order the periodic orbit is the level set $\{E = u^2/2\}$ closed through the short segment at zero. In the conservative-orbit approximation this gives multiplier amplitude $\max \lambda = 2u$, attained at $\ell = \ell^*$. The logit range spans the two roots of $\tau \Delta V(\ell) = u^2/2$ (harmonic limit $\ell^* \pm u/\omega_0$). The period equals the conservative-flow period divided by η, because the steps at zero move at the flow speed at the minimum of the potential, and the correction from the visit to zero cancels at leading order. Visit length is $\sqrt{2\pi/(\eta\omega_0)}$ in the harmonic limit. Measured against direct simulation across a grid of 54 parameter settings in (ζ, τ, η, u), the amplitude ratios fall in $[1.001, 1.029]$ and the logit-range ratios in $[1.002, 1.091]$. Period ratios fall in $[0.998, 1.025]$. The ceiling formula for the visit length matches the simulated length of every recorded visit to zero across six parameter settings.*

B Experimental Setup

B.1 Compute environment

All simulations run on CPU in JAX with double precision enabled, with a reference NumPy implementation for the classifier and the game draws. Symbolic checks use sympy and the exploitability linear programs use scipy. A cell is one choice of the parameters (ρ, ζ, τ), run as its own ensemble of games. A 20,000-game classifier cell takes roughly 2 core-hours, and a 500-game tangent-map-estimate cell at $T = 3 \times 10^5$ roughly 6; the 10^7-step run takes a few hundred, and the whole record on the order of a few thousand. Every cell is deterministic given its integer seed. Per-cell seeds derive from a recorded base seed as base $+ 1000 \cdot$ cell index. Game draws, initializations, and the seeds of the tangent vectors used by the tangent-map estimate of Appendix C all derive from the per-cell seed.

B.2 Default hyperparameters

Unless stated otherwise, $K = 8$ actions, step size $\eta = 0.05$, horizon $T = 30{,}000$, and multiplier upper bound $\lambda_{\max} = 50$. The regime diagram, the fine ζ sweep, and the damping sweep use $G = 20{,}000$ games per cell. Everything else uses $G = 2{,}000$. Strategies initialize with logits i.i.d. $\mathcal{N}(0, 0.1^2)$ and multipliers at zero. Long runs for the tangent-map estimate use $T = 3 \times 10^5$ with 400 games per cell and $w = 4$ tangent vectors (Appendix C). The coarse grid spans $\rho \in \{-1, -0.5, 0, 0.5, 1\}$, $\zeta \in \{0.5, 0.8, 1.2, \infty\}$, and $\tau \in \{0.1, 1, 10\}$. At $\rho = 0$ the fine ζ sweep spans ζ from 0.5 to 1.2 in steps of 0.05, then 1.3 and ∞, at the same three τ. At $\rho = 0$, $\tau = 1$ the damping sweep spans the damping rate δ of the damped update below, $\delta \in \{0, 0.003, 0.01, 0.03, 0.1, 0.3, 1\}$ at seven budget ratios.

B.3 Game generation

For each game two independent $K \times K$ standard Gaussian matrices Z, Z' are drawn. The payoffs are $A = Z$ and $B = \rho Z + \sqrt{1 - \rho^2}\, Z'$ entrywise, indexed by the same joint action. This gives

(2) exactly. $\rho = -1$ makes $A + B = 0$. Each player's cost vector has entries i.i.d. uniform on $[0.5, 1.5]$. The budget is $d = \zeta \cdot \bar{c}$ with $\bar{c}$ the uniform-play cost (the mean entry of the cost vector), as in (3). Because the random-stream order does not depend on ρ, the draws Z, Z' are shared across payoff-correlation rows. Appendix D.9 instead builds always-feasible budgets. Each budget is $d = c_{\min} + s\,(\bar{c} - c_{\min})$, the cheapest action's cost plus a fraction s of the gap to the uniform-play cost, where s is the budget position. Only the budget position varies.

B.4 The three update rules

Let $\mathrm{viol}_x = x_t^\top c_x - d_x$ be player 1's violation (signed) and $\mathrm{MW}(x, g)_k \propto x_k \exp(g_k)$ the multiplicative-weights step.

1. Multiplier-first update: $\lambda_x^{(t+1)} = \Pi_{[0,\lambda_{\max}]}\big(\lambda_x^{(t)} + \tau\eta\,\mathrm{viol}_x\big)$, then $x_{t+1} = \mathrm{MW}\big(x_t, \eta(Ay_t - \lambda_x^{(t+1)} c_x)\big)$, and symmetrically for player 2. This is (4). Both players read the state at time t, so the two of them step together; the ordering that the simultaneous update changes is the one between a player's own multiplier and strategy.
2. Simultaneous update (also analyzed in the theory): the same two maps, but x_{t+1} uses $\lambda_x^{(t)}$ instead of $\lambda_x^{(t+1)}$, so both variables step from the state at time t. On a common set of 500 games the two updates agree in label at 0.91 and their top exponents correlate at 0.91.
3. Damped update: $\lambda_x^{(t+1)} = \Pi_{[0,\lambda_{\max}]}\big(\lambda_x^{(t)} + \tau\eta\,(\mathrm{viol}_x - \delta\,\lambda_x^{(t)})\big)$ with the multiplier-first strategy step. $\delta = 0$ recovers the multiplier-first update. In dimensionless terms the damping range $\delta \in [0.003, 0.03]$ of Section 5.3 corresponds to damping half-lives of 0.4–4 oscillation periods.

Sampled feedback replaces Ay_t by the payoff column of one action drawn from y_t (and symmetrically), with a batch of b draws averaged where stated.

C Evaluation Metrics

C.1 Classifier labels

The classifier labels each game converged, cycling, or separating on the tail (the final 20% of the run). An additional long run records whether the top finite-time exponent exceeds the growth threshold. A game is labeled converged (its strategies stop moving) if the tail-mean per-step strategy movement $\|x_{t+1} - x_t\|_2 + \|y_{t+1} - y_t\|_2$ falls below 10^{-6}. The separation test is a low-cost check for chaos: whether two nearby copies of the same game separate. The test integrates a copy from the same draw with player 1's first logit offset by $\varepsilon = 10^{-8}$. Every 50 steps the logit displacement is renormalized to ε while the multipliers of the copy evolve freely. The separation rate is the mean of $\log(\|\Delta\|_2/\varepsilon)$ over the second-half renormalization events divided by the interval. A non-converged game is separating if this rate exceeds 10^{-3} per step, cycling otherwise. Factor-of-ten shifts of either threshold move the peak of the separating fraction by at most one cell.

C.2 Finite-time Lyapunov exponents

A game is chaotic when the top finite-time Lyapunov exponent Λ_1 of the full state (both players' logits and both multipliers) exceeds the growth threshold under the following tangent-map estimate. We write exponents as Λ throughout, since λ is the multiplier. We propagate $w = 4$ tangent vectors exactly through the linearized multiplier-first update, re-orthogonalize them every 20 steps by a QR factorization, and accumulate the averaged logarithms of the absolute values of the diagonal of R over the second half of the run. An exponent is above the growth threshold when its accumulated log growth exceeds 3, that is, a threshold of $6/T$. The growth threshold is conservative. It produces no positive exponent on an analytic rotation whose exponent is zero in closed form. At the peak cell the tangent-map estimate and the finite-difference estimate give the same chaotic fraction (Appendix D.8). The tangent-map estimate recovers each exponent of an analytic block-rotation reference to 10^{-6}. That reference is the linear map of $\mathbb{R}^4$ whose matrix is block diagonal with two rotation blocks of moduli 1.002 and 0.995 at angles 0.31 and 0.11. Its four exponents are the

logarithms of those moduli, each twice. The tangent-map top exponent correlates at 0.994 with the finite-difference estimate on the same games. Under the growth threshold the chaotic fraction at the peak cell is stable from $T = 3 \times 10^5$ on, whereas a fixed per-step threshold declines with the horizon (Appendix D.1). We therefore quote every main fraction at $T = 3 \times 10^5$ under that threshold.

C.3 Tail violation

The tail violation of a run is the mean of the two players' positive parts, $\frac{1}{2}\big([x^\top c_x - d_x]_+ + [y^\top c_y - d_y]_+\big)$, averaged over the tail. A run in which neither player exceeds its budget contributes zero. Tables that also report a percentile give it as p99 of that per-run quantity across the games of the cell, not across the steps of a run.

C.4 Uncertainty

Fractions come with Wilson score intervals (95%). For $G = 2{,}000$ games per cell the 3σ half-width is at most ± 0.034, wider than the 95% interval. The game draws are the only randomness.

D Additional Experiments

Table A5 maps each claim of Section 5 to the table that records it. The parts below are ordered by topic, not by the rows of that table. Unless a caption or a row says otherwise, every fraction in this appendix is over all the games of its cell, not only the feasible ones. At the peak cell of the fine sweep the chaotic fraction is 0.321 over all games and 0.362 over the feasible ones (Table A21). Appendices D.1 to D.3 give the horizon dependence, the robustness across constructions, and the fractions against per-game slack. The projection learners are there too (Appendix D.3.1). Appendices D.4 to D.6 give what separating games give up, the interventions on the bound at zero, and the full record for sampled feedback and damping. Appendices D.7 to D.10 hold the specifications of the variants, how the exponents and frequencies are measured, the audits of the protocol, and the per-cell records.

D.1 How the chaotic fraction depends on the horizon

The chaotic fraction of Section 5.1 comes from a longer horizon and several exponent estimators (Table A6, Figure A2).

D.1.1 Horizon behavior

About 18% of peak-cell games are chaotic at $T = 30{,}000$. That fraction rises to a third at 3×10^5 and then plateaus through 10^7. Across full-size replications that third is 0.32–0.34. Pooling with the 400- and 800-game cells (0.258 and 0.303) gives 0.319. All three rules for the chaotic label agree here (Table A6; Appendix D.9). Exponents of chaotic games are slow (10^{-5}–10^{-3} per step). The classifier separating fraction at the peak cell is 76% (the classifier is the cheap, high-recall test that selects games for the exponent measurement).

D.1.2 Remainder and robustness

The not-chaotic remainder tends toward zero exponents (marginal exponents shrinking with horizon). Table A25 distinguishes the separating games that are not chaotic from the chaotic ones. An independent 0–1 test confirms the labels (K_{01} medians 0.86 vs 0.01; Appendix D.9). Initialization, burn-in, blockwise estimation, and softplus reparameterization leave the chaotic fraction unchanged. Of the chaotic games 85% persist to 3×10^6 steps. Along the budget ratio the chaotic fraction of the long-run ζ sweep peaks where the classifier separating fraction peaks (classifier recall 0.97 and precision 0.41 at the peak cell; Tables A6, A8 and A23, Figure A2).

D.2 Robustness across constructions

Each variant changes one ingredient of the setup and tests whether the peak of the separating and chaotic fractions of Section 5.1 persists against the matched no-budget cell. Table A36 lists every

Table A1: Separating fraction against the per-game slack m ($\rho = 0$, $\tau = 1$, $K = 8$; seven cells $\zeta = 0.6$ to 1.2, 20,000 games each; feasible games with a converged no-budget run). Columns: games per bin, separating fraction with 3σ interval, share with both multipliers active at the end of the run, and the separating fraction among them. Pooled rows list both bin values and the count.

Slack bin	n	Separating	Both active	Separating given both active
below -0.50	13,901	0.787 [0.777, 0.798]	0.397	0.912
-0.50 to -0.40	7,371	0.766, 0.749	0.366, 0.347	0.923, 0.930
-0.40 to -0.30	7,779	0.744, 0.716	0.338, 0.304	0.924, 0.911
-0.30 to -0.20	7,609	0.708, 0.698	0.301, 0.263	0.904, 0.904
-0.20 to -0.10	6,786	0.676, 0.687	0.242, 0.202	0.887, 0.880
-0.10 to -0.05	3,061	0.684 [0.658, 0.708]	0.160	0.892
-0.05 to 0	2,741	0.593 [0.564, 0.621]	0.077	0.764
0 to $+0.05$	2,542	0.077 [0.062, 0.094]	0.023	0.931
$+0.05$ to $+0.50$	8,717	0.030 to 0.069	at most 0.017	0.80 to 1.00

variant with its verdict under the tangent-map estimate and the classifier, and Table A24 the payoff-correlation rows.

The tested robustness variants retain a peak of the classifier separating fraction at an intermediate budget ratio (mechanism controls in Table 2 in the body; the consolidated verdict table is Table A36; constructions in Appendix D.7). Tested payoff models and update families retain the peak. Heavy-tailed, uniform, sparse, and rank-2 payoffs show it at 0.19–0.29 against ≤ 0.05 with no budget. Projected gradient ascent shows it too at classifier level. Optimism and extragradient also keep the peak across a tenfold step-size sweep at matched $\eta^2 T$ (Table A18).

The peak comes from the constraint structure: two independent budgets amplify it under both the classifier and the tangent-map estimate while a duplicated budget with doubled gain does not (Table A29). Joint-action costs, $K = 32$ actions, $N = 3$ and 5 players, a Markov variant, and every payoff-correlation row except the zero-sum row reproduce it (Tables A18, A30, and A32; Figure 2, right). The zero-sum row has no peak because most of its no-budget runs are already separating.

Step size and horizon enter jointly. At fixed $\eta^2 T$ the chaotic-fraction curves along the budget ratio for a fourfold range of step sizes fall on one curve: the peak position is fixed, the fractions lie within ± 0.05 of the cell mean, and $\Lambda_1 T$ matches cell by cell. Fractions rise only with $\eta^2 T$. Decaying steps stall the peak as predicted, and the choice between the two updates is immaterial (Table A31). Chaos under the tangent-map estimate is established at $K \geq 8$. At $K = 2$ the non-converged population is mostly quasi-periodic, and the chaotic fraction rises smoothly with K (Appendix D.9).

D.3 Separating and chaotic fractions against per-game slack

The separating and chaotic fractions fall as the per-game slack rises (Section 5.1). The peak along ζ is then a product of two monotone shares. No peak lies at slack slightly below zero. The one added variant runs the learner with no budget ($\zeta = \infty$) for each constrained cell on the same drawn games and initial logits with the same seed. We store the tail spending $\bar{x}^\top c_x$ and $\bar{y}^\top c_y$. Per-game slack is $m = \min(m_x, m_y)$ with $m_x = \zeta - \bar{x}^\top c_x / \bar{c}_x$. Here $m < 0$ means the learner with no budget would violate the budget, and a more negative m means a more strongly binding budget. Games are binned by m in steps of 0.05 with 3σ Wilson score intervals. About half of the no-budget runs are not converged at $T = 30{,}000$ (they cycle, as Table A34 records at $\zeta = \infty$). Our primary reading uses the converged half, and a secondary reading keeps every game with its tail-mean spending. The measurement rests on the main tangent-map estimates at the two cells $\zeta = 0.75$ and 0.8 (2,000 games each, $T = 3 \times 10^5$); the classifier on the same 2,000 games across seven values of ζ; the main 20,000-game classifier cells at $\tau = 1$ joined to new no-budget variants; and the main long-run ζ sweep (1,000 games per ζ). Table A1 gives the largest record.

The separating fraction falls monotonically with the slack (Spearman -0.98 over the 22 bins of width 0.05). The fraction is highest and roughly flat where the budget binds strongly, falls by about a fifth as the slack nears zero, and vanishes once the budget no longer binds. The slack acts through the active multipliers. The share of games with both multipliers active falls smoothly from 0.40 far below zero slack to 0.077 just below it and 0.023 just above. Among games with both multipliers active the separating fraction is 0.88 to 0.93 in every populated bin and does not depend on the slack. Three other records agree in shape. On the tangent-map estimates at $\zeta = 0.75$ the chaotic

fraction among feasible games is 0.40 below $m = -0.5$, 0.39 to 0.46 between -0.5 and -0.4, 0.28 to 0.43 between -0.4 and -0.15, 0.15 to 0.30 in the two bins just below zero, and 0.00 to 0.07 above zero. At $\zeta = 0.8$ the same order holds (0.33 to 0.42 below -0.2, 0.17 and 0.07 just below zero, 0.03 to 0.07 above). On the classifier over the same 2,000 games across seven values of ζ the pooled feasible curve is 0.81 below -0.5, 0.75 to 0.78 from -0.5 to -0.3, 0.68 to 0.72 from -0.3 to -0.05, 0.57 just below zero, and 0.07 just above (Spearman -0.94). Keeping every game whether or not its no-budget run is converged gives 0.81, 0.78, 0.69 to 0.74, 0.60, and 0.17 in the same bins. On the long-run ζ sweep (1,000 games per ζ) the feasible curve is 0.37 to 0.39 below -0.35, 0.29 to 0.31 from -0.35 to -0.15, 0.19 in the two bins from -0.15 to -0.05, 0.09 just below zero, and 0.02 to 0.05 above. Games whose no-budget run is not converged have chaotic fractions 0.377 and 0.336 in the two cells $\zeta = 0.75$ and 0.8, the same as the converged runs. We fixed two hypotheses before the runs. A peak at slack slightly below zero would show as a maximum within 0.1 of zero slack with lower values on both sides. No record shows one, as every maximum is at $m \leq -0.35$. The bins just below zero lie below the strongly binding bins, beyond their 3σ intervals (0.593 against 0.787 in Table A1). A feasibility transition is a monotone rise as the slack turns more negative with no interior maximum. The rise holds on both classifier records (Spearman -0.98 and -0.94). Both tangent-map records put their maximum one bin inside a flat plateau (Spearman -0.90 and -0.75). The plateau is monotone within 3σ. Along ζ the median slack moves from -0.54 at $\zeta = 0.6$ to $+0.05$ at 1.2 while the feasible fraction rises from 0.230 to 1.000. Among feasible games the share whose budget binds falls from 0.99 at $\zeta = 0.6$ to 0.39 at 1.2. At the peak cell the separating fraction among feasible games is 0.81 when the budget binds and 0.16 when it does not bind. The separating fraction over all games therefore comes from a product: feasible fraction times binding share. That product peaks at $\zeta = 0.8$. The peak of the separating fraction over all games lies one grid step below at 0.75 because infeasible games still add 0.071 there (Table A20). Appendix D.9 treats the always-feasible budgets the same way.

D.3.1 Feasible-set projection learners

Two learners keep the strategy inside the feasible set and learn no multiplier. They run the same feasible games, to test whether separating and chaotic runs need the multiplier at all (matched games, seeds, and initial strategies; $\rho = 0$, $K = 8$, $G = 2{,}000$, $T = 30{,}000$, $\eta = 0.05$; Table A2). The Euclidean learner projects a gradient step in Euclidean distance onto the simplex intersected with the budget half-space. The KL learner takes the multiplicative-weights step and applies the KL projection onto the budget half-space. This step multiplies the strategy by $e^{-\mu c}$ with the smallest $\mu \geq 0$ that meets the budget. The KL learner keeps the primal geometry of the primal-dual learner and replaces only the learned multiplier by a projection multiplier. Both are separating in at most 0.032 of feasible games at every budget ratio where the primal-dual learner is separating in 0.28 to 0.82, agree with each other within 0.04 on every cell, and never violate a budget. In a long run at $\zeta = 0.75$ ($G = 1{,}000$, $T = 3 \times 10^5$) the KL learner is chaotic on 0.012 of games against 0.321 for the primal-dual learner on the same seed. The tangent-map estimate uses the implicit derivative of the projection multiplier. The Euclidean learner gives 0.006 with the finite-difference estimate. Both are converged on about 0.95 of feasible games at $\zeta = 0.6$ but only on 0.75 at $\zeta = 0.75$ and 0.5 at $\zeta = 1.2$. The remainder cycle with a negative exponent, as projected gradient with no budget does (0.48 converged and 0.49 cycling at $\zeta = \infty$). The last column is projected gradient with a learned multiplier on the same feasible games: it is separating in 0.05 to 0.15 of them, at least three times that of the Euclidean projection learner at every budget ratio from 0.7 to 1.0. Within one primal geometry the learned multiplier is what makes runs separating.

D.4 What separating games give up

We measure exploitability and tail violations because a non-convergence label does not measure welfare (Appendix D.8). These measurements show what the classifier labels of Section 4 mean. Separating games at $\zeta = 0.7$ to 0.8 have median last-iterate exploitability of 0.27–0.32 (Table A14). Only 6 games of 4,050 jointly reach a small exploitability and a small tail violation. Time-averaging splits the labels: the exploitability of cycling games falls near $1/T$ while that of chaotic games barely decays (Appendix D.9). The last iterate is far from equilibrium. The averaged marginals are slow to reach Nash. Under own-strategy costs the empirical joint distribution is near a coarse correlated equilibrium under feasible deviations (Appendix D.8).

Table A2: Feasible-set projection learners without a multiplier against the primal-dual learner on the same feasible games at $\rho = 0$, $\tau = 1$, $K = 8$, $G = 2{,}000$ per cell: converged / cycling / separating fractions (cycling = neither converged nor separating; the three sum to one within rounding). The last column is projected gradient with a learned multiplier on the same feasible games.

ζ	primal-dual (multiplicative weights)	Euclidean projection	KL projection	projected gradient with multiplier
0.6	0.124 / 0.056 / 0.820	0.946 / 0.054 / 0.000	0.946 / 0.054 / 0.000	0.046 / 0.902 / 0.051
0.7	0.102 / 0.101 / 0.797	0.822 / 0.175 / 0.003	0.826 / 0.173 / 0.001	0.056 / 0.867 / 0.078
0.75	0.098 / 0.131 / 0.770	0.754 / 0.243 / 0.003	0.758 / 0.240 / 0.001	0.065 / 0.839 / 0.097
0.8	0.094 / 0.166 / 0.739	0.685 / 0.306 / 0.009	0.698 / 0.294 / 0.007	0.072 / 0.811 / 0.117
0.9	0.117 / 0.238 / 0.645	0.555 / 0.424 / 0.022	0.589 / 0.394 / 0.018	0.111 / 0.740 / 0.150
1.0	0.168 / 0.301 / 0.531	0.516 / 0.461 / 0.023	0.544 / 0.435 / 0.021	0.155 / 0.726 / 0.119
1.2	0.301 / 0.417 / 0.282	0.488 / 0.480 / 0.032	0.532 / 0.438 / 0.030	0.275 / 0.667 / 0.058

D.5 Intervening on the bound at zero

The necessary-conditions paragraph of Section 5.2 rests on changing the bound at zero, the projection, or the multiplier's integral of past violations one at a time (Tables A13 and A18).

The peak needs the multiplier's integral of past violations and a reachable low-multiplier region; the kink is not needed. Two smooth interventions on the same games retain the peak (Table A12, Appendix D.7).

Removing the integral of past violations removes the peak and changing the feedback gain does not. The penalty variant and sliding-window multipliers with short windows remove the peak, and a raised lower bound barely moves it. A sliding-window multiplier recovers the peak once its window reaches the oscillation period. The penalty-variant and raised-lower-bound readings hold under both the classifier and the tangent-map estimate (Tables A13 and A18), and the sliding-window reading under the classifier (Table A13). Mean multipliers fall both in the sliding-window multipliers with short windows and in the $W = 1000$ variant that keeps the peak. The fall of the multiplier is therefore not what removes the peak. Oscillations are less frequent in the variants that remove the peak (share of games with a dominant oscillation peak 0.20–0.26 against 0.51–0.56, Appendix D.9).

The leading Lyapunov vector and the growth lie in different directions. The squared norm of the leading Lyapunov vector concentrates on the directions of low-probability actions (0.87 squared-norm share at the peak cell, Table A6), while the growth is generated in the oscillator directions. The cost-direction-only strategy update collapses the chaotic fraction from 0.31 to 0.034, and the two-action subgame collapses it from 0.32 to 0.056, each against the baseline of its own run. Both interventions also remove the oscillation itself (Appendix D.8). The projected copy distinguishes the two sets of directions: confining the deviation to the oscillator directions gives a chaotic fraction of 0.275 against 0.319 unconfined. Confinement to the other directions, multipliers kept, gives 0.054 (Table A22). A projected tangent vector kept off the oscillator directions stops growing, while a projected tangent vector confined to them grows at full scale (Appendix D.8). Expansion precedes visits to the low-multiplier region (99%; Appendix D.8).

With one constrained player, player 2 runs unconstrained multiplicative weights with no budget and no multiplier while player 1 runs the primal-dual learner. The draws match the run with both players constrained (Table A3, games feasible for player 1). The variant tests whether separating and chaotic runs need two constrained players. At $\zeta = 0.75$ the games with one constrained player are separating in 0.537 against 0.770 with both constrained. In the $\zeta = 0.75$ long run ($G = 1{,}000$, $T = 3 \times 10^5$, tangent-map estimate) the chaotic fraction is 0.173 of all games and 0.184 of those feasible for player 1, against 0.321. One oscillator coupled to an adapting opponent therefore accounts for most of the separating and chaotic fractions. With both constrained, one active multiplier is the common configuration: among feasible games exactly one multiplier is active at the end of the run in 0.43 to 0.54 across $\zeta = 0.6$ to 1.0 (Table A21).

D.6 Sampled feedback and damping: the full record

The sampling and damping claims of Section 5.3 rest on replacing exact feedback by sampled payoffs and costs and on damping the multiplier (Tables A15, A16 and A26).

Table A3: One constrained player against two (same draws, $\rho = 0$, $\tau = 1$, $K = 8$, $G = 2{,}000$, $T = 30{,}000$; player 2 has no budget and no multiplier). Columns: converged / cycling / separating fractions among games feasible for the constrained players, and the share of one-constrained games with an active tail multiplier. Long run at $\zeta = 0.75$: 0.173 chaotic against 0.321.

ζ	one constrained: conv. / cyc. / sep.	multiplier active	two constrained: conv. / cyc. / sep.
0.6	0.280 / 0.151 / 0.570	0.828	0.124 / 0.056 / 0.820
0.75	0.239 / 0.224 / 0.537	0.730	0.098 / 0.131 / 0.770
0.9	0.257 / 0.310 / 0.433	0.608	0.117 / 0.238 / 0.645
1.2	0.395 / 0.445 / 0.161	0.255	0.301 / 0.417 / 0.282

Table A4: Perturbation-scale sweep at the peak cell: fraction of all games whose two copies are more than 10^{-3} apart at T, by variant, batch size b, and initial displacement ϵ (common random numbers, no renormalization).

variant	b	$\epsilon = 10^{-8}$	10^{-6}	10^{-4}	10^{-2}
exact	1	0.154	0.390	0.939	1.000
Gaussian noise	1	0.359	0.652	0.953	1.000
Gaussian noise	16	0.204	0.490	0.942	1.000
Gaussian noise	256	0.167	0.403	0.936	1.000
sampled	1	0.000	0.001	0.073	1.000
sampled	16	0.001	0.013	0.312	1.000
sampled	256	0.003	0.130	0.821	1.000
sampled	1024	0.011	n/a	n/a	1.000

Under sampled feedback we distinguish deterministic chaos, stochastic non-convergence, and noise synchronization (two copies driven by common random numbers converging to each other). Under batch averaging the separation re-emerges faintly by $b = 4096$ (Appendix D.9). A perturbation-scale sweep at the peak cell ($\zeta = 0.75$, $G = 2{,}000$, $T = 30{,}000$) runs two copies of the learner with sampled payoff and cost feedback under common random numbers. Initial logits differ by ϵ on one coordinate, with no renormalization. The comparison is against the learner with exact feedback and against a smooth control that adds Gaussian noise to the exact fields, with matched per-coordinate variance and common random numbers (Table A4). The distance between two sampled-feedback copies is unchanged until the first step at which the two copies draw a different sampled action, because the sampled action indices are locally constant in the state between such steps. The fraction of games in which the copies ever disagree tracks the fraction of games whose copies end apart (0.075 against 0.073 at batch one and $\epsilon = 10^{-4}$). At $\epsilon = 10^{-2}$ the copies of the sampled variant end apart in every game, and at that scale sampling removes only the infinitesimal signature. At $\epsilon \leq 10^{-4}$ the suppression is set by the probability that the perturbation flips a sampled index. That probability grows with the batch size.

A shorter damping half-life removes the peak (Table A26); the range of Section 5.3 corresponds to half-lives of 0.4–4 oscillation periods (Appendix B).

D.7 Specifications of the robustness-set variants

These constructions back the robustness verdicts of Section 5.1 and the mechanism controls of Section 5.2 (Tables 2 and A36).

D.7.1 Shared conventions

All robustness-set and stabilization cells run at $\rho = 0$. Cost vectors are uniform on $[0.5, 1.5]$ in every variant.

D.7.2 Update-rule variants

The optimistic family applies past-gradient extrapolation to both blocks and keeps the multiplier-first update of Section 4. With $v^{(t)}$ the violation at the current strategies, $g^{(t)}$ the Lagrangian payoff gradient evaluated at the freshly updated multiplier, and logits ℓ,

$$\lambda^{(t+1)} = \Pi_{[0,\lambda_{\max}]}\big(\lambda^{(t)} + \tau\eta\,(2v^{(t)} - v^{(t-1)})\big), \qquad \ell^{(t+1)} = \ell^{(t)} + \eta\,(2g^{(t)} - g^{(t-1)}), \tag{82}$$

Table A5: Claim-to-record index: the table of this appendix behind each main-text claim.

Main-text claim	Record
Outcome fractions over the full grid	Table A34; fine ζ sweep and damping sweep in the paper's records
Feasibility split of the separating-fraction peak	Table A20
Second-level split by active multipliers	Appendix D.9 (feasibility, budget positions, and the feasibility split)
Always-feasible budgets, tangent-map estimate and classifier	Table A19
Separating fraction against the per-game slack	Table A1
Cross-distribution peak positions	Appendix D.9 (text)
Long-run ζ sweep and the classifier checked against it	Table A8, Table A23, Table A25; games that are not chaotic: Appendix D.1 (text)
Stability of the chaotic fraction	Table A6
Estimator and convention audits	Table A6
Verdicts on the robustness-set variants	Table 2 (body, mechanism controls); Table A36 (consolidated)
Interventions at zero and on the multiplier's integral of past violations (classifier)	Tables A12, A13
Intervention and variant Lyapunov spectra (tangent-map estimate)	Table A18
Location of the expansion	Table A6; Table A9
Frequency formula: strata and selection	Table A10; Table A7
Exploitability and the four-way split	Table A14; Table A17
Sampled feedback	Table A15
Damping tradeoff and tail-violation percentiles	Table A26; Figure A5
Protocol audits	Table A6 and the tables of Appendix D.9

with $g^{(-1)} = g^{(0)}$ and $v^{(-1)} = v^{(0)}$. The first step is then the plain step. Our extragradient family is the simultaneous mirror-prox form (extragradient in the mirror geometry): a half step from $(\ell^{(t)}, \lambda^{(t)})$ with the field evaluated there,

$$\begin{aligned} \bar{\lambda} &= \Pi_{[0,\lambda_{\max}]}\big(\lambda^{(t)} + \tau\eta\, v(x_t)\big), & \bar{\ell} &= \ell^{(t)} + \eta\, g(x_t, \lambda^{(t)}), \\ \lambda^{(t+1)} &= \Pi_{[0,\lambda_{\max}]}\big(\lambda^{(t)} + \tau\eta\, v(\bar{x})\big), & \ell^{(t+1)} &= \ell^{(t)} + \eta\, g(\bar{x}, \bar{\lambda}), \end{aligned} \tag{83}$$

so the full step uses the half-point violation for the multiplier and the half-point strategies and multiplier for the payoff gradient, with $\bar{x}$ the strategy of the logits $\bar{\ell}$. Both new update families converge last-iterate on a reference game. The check is unconstrained matching pennies (the two-action zero-sum game in which one player wins on matching actions). Plain ascent cycles there.

D.7.3 Payoff-distribution variants

The payoff-distribution variants draw independent entries for the two players, standardized to zero mean and unit variance: Student-t_3 entries scaled by $1/\sqrt{3}$; uniform entries on $[-\sqrt{3}, \sqrt{3}]$; sparse entries as a standard Gaussian times an independent Bernoulli(0.25) mask per entry and per player, scaled by $1/\sqrt{0.25}$; and rank-2 payoffs $A = UV^\top/\sqrt{2}$ with i.i.d. standard Gaussian factors $U, V \in \mathbb{R}^{K\times 2}$ drawn independently per player.

D.7.4 Order, step-size, and geometry variants

The simultaneous-update variant changes only the sequencing of Eq. (4): both blocks step from the state at time t, and the strategy update uses the pre-update multiplier. Both updates therefore coincide whenever the multipliers stay at zero. The variant is bitwise the multiplier-first update on no-budget cells and on a hand-stepped constrained update. Tangent-map estimates of the Lyapunov spectra on the same 500 drawn games under both updates give chaotic fractions 0.310 (multiplier-first) vs 0.282 (simultaneous). Per-game chaotic-label agreement is 0.912, the top-exponent Pearson correlation 0.909 (Spearman 0.753), and the Jaccard overlap of the sets of chaotic games 0.741. Step-size variants run plain ascent at $\eta = 0.1$ and $\eta = 0.02$ with all other protocol constants unchanged. Each runs at the standard horizon $T = 30{,}000$ and at a horizon holding $\eta^2 T$ at its standard value ($T = 7{,}500$ and $T = 187{,}500$ respectively). The projected-gradient variant replaces the mirror step by $x_{t+1} = \Pi_{\mathcal{S}_K}(x_t + \eta\, g_t)$ with the Euclidean projection onto the simplex. The

update stays multiplier-first, and the classifier perturbs, renormalizes, and re-projects in strategy space. The peak test, set before the runs, is evaluated in both the separating-fraction and the non-converged-fraction ($1 - \mathrm{conv}$) forms.

D.7.5 CONSTRAINT-CONSTRUCTION VARIANTS

The variant with two independent budgets draws a second cost vector i.i.d. uniform on $[0.5, 1.5]$, independent of the first, with its own budget ζ times its mean and its own multiplier coordinate. The Lagrangian payoff gradient subtracts $\sum_m \lambda_m c_m$ and both dual coordinates follow the same projected ascent. The joint-action-cost variant replaces each player's cost vector by a matrix with i.i.d. $U(0.5, 1.5)$ entries over joint actions. Player 1's violation is $x^\top C_x y - \zeta\, \overline{C_x}$ with Lagrangian payoff gradient $Ay - \lambda_x(C_x y)$, and symmetrically for player 2 through $C_y^\top x$. With row-constant matrices the variant reproduces the single-budget primal-dual learner (one step to 3×10^{-17}, full-cell outcome fractions within 0.005). In the duplicated-budget variant the second vector is a copy of the first. This keeps the two multipliers identical. The duplicated budget equals a single constraint with doubled dual gain and doubled upper bound. Exact feasibility for the conditioned comparison is the cheapest-action test in the single-budget and duplicated-budget variants. With two independent budgets the test is a linear program over the simplex. In the N-player variants, player μ's payoff is an N-index Gaussian tensor (own action first, remaining players in cyclic order). Entries at the same joint action tuple are jointly Gaussian across players with pairwise correlation $\rho/(N-1)$. Distinct tuples are independent. At $N = 2$ this reduces to Eq. (2), and in the zero-sum row the entries sum to zero almost surely. Entry standard deviation is $K^{(N-2)/2}$, chosen so that the payoff gradient under uniform play has standard deviation $K^{-1/2}$ for every N. This keeps η, T, and the classifier thresholds comparable across the player axis. Each of the N players has one linear cost constraint on its own strategy, with costs, budget, and its own multiplier exactly as in the two-player setting.

D.7.6 INTERVENTION VARIANTS AT ZERO

The interventions at zero keep the strategy update of Eq. (4) and replace only the multiplier step. The softplus-smoothed projection is $\lambda_{t+1} = \min\big(\mathrm{softplus}_\beta(\lambda_t + \tau\eta\, v_t),\, \lambda_{\max}\big)$ with $\mathrm{softplus}_\beta(u) = \log(1 + e^{\beta u})/\beta$. This recovers the hard projection as $\beta \to \infty$ and rounds the kink over a width of order $1/\beta$. The variant with the C^1 lower-bound map replaces the projection by $s_\delta(u) = u$ for $u \geq \delta$ and $\delta\, e^{(u-\delta)/\delta}$ below, a C^1 map with slope one at the junction. This keeps the λ-space gain exactly one away from a width-δ neighborhood of zero. The softplus reparameterization stores z_t with $\lambda = \mathrm{softplus}(z)$: $z_{t+1} = \min(z_t + \tau\eta\, v_t,\, z_{\max})$ with $z_0 = \mathrm{softplus}^{-1}(0.01)$ and $z_{\max} = \mathrm{softplus}^{-1}(\lambda_{\max})$. This map is smooth everywhere and λ stays positive. The C^1 lower-bound map isolates the gain, and the sliding-window multiplier the length of the integral of past violations.

The variant with a raised lower bound projects the multipliers onto $[\lambda_{\min}, \lambda_{\max}]$ with $\lambda_{\min} \in \{0, 0.01, 0.03, 0.1, 0.3, 1\}$ and initializes them at $\lambda_{\min}$. At $\lambda_{\min} = 0$ it is bitwise the primal-dual learner (exact at one and 500 steps). The penalty variant removes the multiplier state entirely: at every step $\lambda = \Pi_{[0,\lambda_{\max}]}(\kappa\,[v]_+)$ is recomputed from the current violation with $\kappa \in \{5, 20, 80\}$. The penalty variant then has the same bound at zero and gain as the projection but no integral of past violations: it is recomputed from the current state each step, and $\kappa = 0$ is bitwise unconstrained ascent. The sliding-window multiplier sets $\lambda_t = \Pi_{[0,\lambda_{\max}]}(\tau\eta \sum_{s=t-W+1}^{t} v_s)$ via a ring buffer, the length-W truncation of the unprojected integral at matched ramp gain. $W = 1$ is bitwise the penalty variant, and the large-W limit approaches the multiplier's integral of past violations up to the difference between projecting the output and projecting the state. No multiplier computed from finitely many past violations can enforce exact complementarity (zero mean violation whenever the multiplier is positive), because its fixed point requires a nonzero mean violation $\lambda^*/(\tau\eta W)$. All intervention variants share their drawn games and seeds with the matched primal-dual learner. It reproduces the cells of the fine ζ sweep within the $G = 2{,}000$ sampling interval.

D.7.7 MARKOV GAME VARIANT

The constrained Markov game variant has $S = 4$ states and $K = 4$ actions per player. Per-state payoff pairs are drawn as in Eq. (2), and per-state costs are i.i.d. uniform on $[0.5, 1.5]$. Transition kernels $P(s' \mid s, a_1, a_2)$ are i.i.d. Dirichlet$(1, \ldots, 1)$ over next states, the discount is 0.9, and the

initial distribution is uniform. The cost value is normalized by one minus the discount, and each budget is ζ times the exact uniform-play cost value. The violation reduces at $S = 1$ to the one-shot violation. Each step updates the multipliers first on the exact violation. The step then adds the action-mean-centered Lagrangian Q-values from exact policy evaluation to the logits of every state, the softmax natural-gradient analog of Eq. (4). At $S = 1$ the induced strategy dynamics reduces exactly to the one-shot update. Our classifier is unchanged with the step norm summed over states. The policy evaluation agrees with a direct linear solve and a Monte-Carlo rollout, and the $S = 1$ variant reproduces the one-shot outcome fractions on same-seed games.

D.7.8 SAMPLED-FEEDBACK VARIANTS

The sampling-noise variant replaces the exact fields by unbiased one-sample estimates. Each player observes the payoff vector against one opponent action drawn from the opponent's current strategy, and the multiplier ascends on one sampled own cost minus the budget. Actions are drawn through per-step common random numbers and the inverse cumulative distribution function (CDF) of each trajectory. The sampled payoff and cost fields are unbiased for those of Eq. (4); the projected update of the state is not linear in them. The separation test couples two trajectories through common random numbers. Two variants with a single sampled feedback run on the same seeds. One samples only the payoff feedback while the multiplier ascends on the exact own spending, and the other samples only the cost feedback of the multiplier while the strategy sees the exact field. Batch variants average b such samples per field per step with b up to 4096 (variance scales as $1/b$; a counting reformulation is used from $b = 64$). The independent-noise control drives the separation test's copy with its own uniform stream.

D.7.9 FEASIBLE-DEVIATION REGRET

At each averaging horizon the feasible-deviation regret evaluates each player's feasible best response against the opponent's prefix-averaged marginal minus the running mean of the realized expected payoff $x_t^\top A y_t$, with positive parts summed. This is the feasible-deviation regret of the empirical joint distribution, paired with the violation of the averaged marginals. Here the empirical joint distribution is the running average of the product measures $x_t \otimes y_t$ (the expectation level, before any action is sampled). This feasible-deviation notion restricts each player's deviations to its own budget set. When all mass is feasible this statistic coincides with the coarse correlated equilibrium of the constrained game, with deviations ranging over the feasible simplex. When play puts mass on infeasible strategies the statistic remains well-defined as a feasible-deviation regret. We therefore report the feasible-deviation regret together with the violation column.

D.8 HOW THE EXPONENTS AND FREQUENCIES ARE MEASURED

The frequency-formula test and the exponent measurements of Section 5.2 rest on varying or auditing the measurement procedure (the exponent estimator, its conventions, and the frequency pipeline) while the games and seeds stay fixed. Table A10 shows that the median ratio does not depend on the inclusion rule while the slope does. Keeping only the $2{,}345$ oscillators in the fixed resolvable range sharpens the pooled slope to 0.92 and the median ratio to 1.02. By step-size ratio, the median ratio of measured to predicted frequency over the included oscillators is 1.53 at $\tau = 0.1$ (521 oscillators), 1.03 at $\tau = 1$ (872 oscillators), and 1.02 at $\tau = 10$ ($1{,}408$ oscillators), as the right panel of Figure 4 shows.

D.8.1 ESTIMATOR CONVENTION

Every Lyapunov-spectrum number at $T \geq 3 \times 10^5$ for a differentiable map is a tangent-map estimate ($w = 4$ tangent vectors unless stated, QR every 20 steps, $T/2$ burn-in, growth threshold $6/T$, audited derivative convention at the kink). Lyapunov-spectrum numbers at $T = 3 \times 10^4$ are finite-difference estimates, cross-validated against the tangent-map estimate below. All quoted exponents are finite-time Lyapunov exponent estimates.

Table A6: Measurement audits, one row per audit: estimator conventions, the location of the leading Lyapunov vector, timing of the low-multiplier region, and the stability of the chaotic fraction at the peak cell (chaotic fraction 0.32 baseline; same 1,000 games and seeds unless stated). Tangent-map estimate = exact propagation of the tangent map (the linearized map composed along the trajectory).

Audit	Measurement
Estimator	
Skew-tent and block-rotation analytic references	exponents recovered to 3×10^{-5} and 10^{-6}, finite-difference and tangent-map estimates alike
Offset $\varepsilon \in \{10^{-6}, 10^{-8}\}$; 10^{-10}	medians move $< 8\%$, chaotic fractions unchanged; at 10^{-10} medians inflate 4–8× (float64 precision limit of the finite-difference estimate)
Reorthonormalization interval $\{5, 20, 80\}$	chaotic fractions unchanged
Finite-difference vs tangent-map estimate, same games, $T = 3 \times 10^4$	top-exponent correlation 0.994; per-game median ratio within 2%; identical positive fractions
Number of tangent vectors $w = 1$ / 8 / 18 vs $w = 4$	label agreement 0.97 / 0.985 / 0.99; top-exponent correlations 0.983 / 0.994 / 0.992
Derivative convention at the kink (tie derivative 0.5 / 1 / 0)	bit-identical Lyapunov spectra; 0 exact ties at the kink in 6×10^7 updates
Leading Lyapunov vector, chaotic games at $\zeta = 0.75$ / 0.8 ($T = 3 \times 10^4$, finite difference, own 1,000-game cells; $n = 176$ / 202 chaotic; medians)	
Squared projection onto the oscillator directions (cost-gradient direction $J_{\mathrm{softmax}}^{\top} c$)	0.024 / 0.024 (random-direction baseline $4/18 = 0.22$)
Conditional-to-full top-exponent ratio (multiplier tangents zeroed)	1.51 / 1.30
Squared-norm shares (p1 logits / p2 logits / multipliers)	0.51 / 0.48 / 0.001 (both cells)
Squared-norm share by action probability (low-probability < 0.01 / middle / support > 0.1 / multipliers)	0.87 / 0.02 / 0.10 / 0.001 (both cells)
Timing of the low-multiplier region (growth difference = growth of the leading Lyapunov vector in low-multiplier periods minus that in other periods)	
Growth difference at $\zeta = 0.75$ / 0.8 (threshold 0.5)	-0.004 / -0.002; correlation -0.03 to -0.02; low-multiplier time fraction 0.48–0.62
Threshold sweep 0.25 / 0.5 / 1.0	growth difference within -0.005 to -0.002
Lagged cross-correlation, growth leads / zero lag / growth follows by 1–3 intervals	-0.01 to $+0.01$ / -0.02 / -0.06 to -0.09
Stability of the chaotic fraction at the peak cell	
Five independent strategy initializations	fractions 0.317–0.322; per-game agreement 0.94
Initialization scale 0.1–0.5 and $\lambda_0 \sim U[0, 2]$ (classifier, three cells around the peak)	peak moves ≤ 0.03 ($0.756 \to 0.729$–0.755)
Burn-in $T/4$ / $T/2$ / $2T/3$	0.343 / 0.320 / 0.292
Six post-warm 25,000-step blocks (chaotic games)	all six positive in 75%, at least five in 89%; block bootstrap 0.32 $[0.31, 0.37]$
Block confidence-interval certification vs growth threshold	0.316 vs 0.320; per-game differences symmetric
Horizon sequence $3 \times 10^5 \to 10^6 \to 3 \times 10^6$	291 stay / 30 drop / 37 enter, then 301 / 27 / 21: $\sim$ 9% churn per tripling, net flat; 85% of the 321 chaotic games persist through both longer horizons
Step-size limit ($\eta = 0.02$, $\eta^2 T$ fixed)	0.327 vs 0.321, per-game agreement 0.93; exponents scale by 0.38 $\approx$ the η ratio

D.8.2 Lyapunov-spectrum estimator and validation

The finite-difference estimate of the Lyapunov spectrum is Benettin/QR over the full state (both players' logits and both multipliers). It uses $w = 4$ perturbed copies of the full nonlinear trajectory at offset $\varepsilon = 10^{-8}$ along an orthonormal random frame. The displacement matrix is QR-reorthonormalized every 20 steps and exponents are accumulated as the averaged log of the R diagonal over the second half of the run. Table A6 collects the estimator audits: analytic-reference recovery, operating-point invariance, the cross-check of the finite-difference estimate against the tangent-map estimate, and the scan over the number of tangent vectors. Finite-difference estimates agree with tangent-map estimates to the precision the table states.

A second zero-exponent direction besides the softmax shift direction would occupy the top of the Lyapunov spectrum, because QR orders the exponents only up to the finite-horizon resolution interval. At $T = 3 \times 10^5$ the resolution interval $\pm 2 \times 10^{-5}$ contains both leading exponents of every game that is not chaotic, so the not-chaotic remainder is neutral at that resolution (Appendix D.1). The initial frame is an orthonormalized standard-Gaussian draw from a fixed seed stream (stated per run). Propagation is in logit coordinates, where the softmax shift direction (a common shift of one player's logits) is neutral. It contributes one zero direction per player (the exponent of the all-ones logit direction, identified by its exact zero growth). That direction is counted and not treated as quasi-periodicity.

D.8.3 Estimator policy at long horizons

The horizon sequence below is the same set of cells rerun at a sequence of increasing horizons. At long horizons the two estimators diverge, and the tangent-map estimate is authoritative. A same-cell control stated before the runs at $T = 200{,}000$ finds rank correlation 0.96 but an additive finite-difference excess concentrated in the low-growth strata. Agreement is within 2% for top exponents above 10^{-4}, with an excess of 1–3 $\times 10^{-5}$ per step below. This excess signals the subtraction-noise level of the finite-difference estimate. At $T = 30{,}000$ the growth threshold (2×10^{-4}) is far above this noise level and both estimators agree. At $T = 300{,}000$ the threshold (2×10^{-5}) lies inside it. This is why every long-horizon Lyapunov-spectrum statistic in this paper is a tangent-map estimate. The supplementary material records the comparison. Block-rotation recovery (10^{-6}; Table A6) and the peak-cell classifier validate the tangent-map estimate.

D.8.4 Location of the expanding direction

The modal-projection measurement propagates the same tangent basis. At each late-run reorthonormalization it projects the leading Lyapunov vector onto the four-dimensional subspace spanned by the two multiplier axes and each player's normalized cost-gradient direction in the logits $v_i \propto J_{\mathrm{softmax}}(\ell_i)^\top c_i$. Supports are disjoint, and the basis is orthonormal by construction. At vertex-saturated states the cost-gradient direction degenerates and is set to zero. Recording the projection leaves the Lyapunov spectrum untouched (bit-identical exponents with and without it). Our conditional-exponent measurement propagates a second set of tangent vectors along the same base trajectory with multiplier perturbations zeroed after every tangent step (base flow unchanged; it matches a per-game Jacobian computation). The mode-localization measurement splits the squared-norm share of the leading Lyapunov vector in the strategy block by the current action probabilities. Table A6 collects the measurements. The leading Lyapunov vector lies in strategy space, with its squared norm concentrated on the directions of low-probability actions. A conditional-to-full ratio above one means the multiplier coupling contracts along the leading Lyapunov vector (the folding acts through that coupling). Setting that coupling to zero at the tangent level raises the exponent.

D.8.5 Projected tangent vectors

A sharper pair of projected tangent vectors localizes where growth is generated ($G = 300$, peak cell, 60 blocks over the full run). One tangent vector is re-projected at every renormalization onto the oscillator directions. Per player these directions are the logit direction that moves $x^\top c$ together with the multiplier axis. Among chaotic games that vector grows at full-exponent scale: block-rate median 1.8×10^{-4} against top median 2.4×10^{-4}, positive in 0.84, rank correlation 0.53 with the full exponent. Games that are not chaotic give 1.5×10^{-6}, positive in 0.20. The other tangent vector, re-projected off the oscillator directions, stops growing (positive in 0.05 of chaotic games). The fixed-direction projections of the conditional-exponent measurement below remove only the constant cost vector of the game and retain growth.

D.8.6 Two-action subgame control

Each peak-cell game is restricted to two actions per player: the cheapest and the strongest against a uniform opponent, with the second-strongest substituted when the two coincide. The rule was fixed before the runs. Budgets are renormalized to the same budget ratio ζ over the retained actions, a matched-budget-ratio control. Its cost rescale (median 1.20) enters the multiplier gain. That restriction collapses the chaotic fraction from 0.32 to 0.056, the $K \leq 3$ family level. The cost-

direction-only strategy update, the conditional Lyapunov exponent, and the two-action subgame agree. An original-budget variant keeps the per-game budgets d and leaves the costs unscaled. Its per-game budget ratio has median 0.91 and deciles 0.76–1.13, about a quarter of them above one. Its chaotic fraction is 0.13, a retention that tracks its heterogeneous binding.

D.8.7 Derivative convention at the kink

At the kink of the projection the three branch-derivative conventions give bit-identical Lyapunov spectra (Table A6). On the branch at zero or at the upper bound the pre-projection input is strictly outside the boundary (0 events within 10^{-12} of a tie in 6×10^7 updates), so the convention is immaterial.

D.8.8 Softplus Lyapunov spectrum and low-multiplier timing

The softplus reparameterization is smooth everywhere below the upper bound and has no projection. The softplus Lyapunov spectrum applies the same tangent-map estimate to it. Chaotic fractions are 0.064, 0.288, 0.297, 0.127 at $\zeta = 0.6$, 0.75, 0.9, 1.2 ($T = 120{,}000$, growth threshold 5×10^{-5}). This interior peak numerically matches that of the primal-dual learner at the same horizon (0.292 at the peak cell): the chaotic-fraction peak needs no kink. This match extends beyond fractions to magnitudes at that horizon. Chaotic-game top exponents have median 1.4×10^{-4} per step under softplus against 2.0×10^{-4} under the hard projection, with overlapping interquartile ranges ($[0.8, 2.3] \times 10^{-4}$ vs $[0.7, 4.4] \times 10^{-4}$; softplus at $T = 1.2 \times 10^5$). At the matched $T = 3 \times 10^5$ horizon the fractions still agree (softplus 0.362 at $G = 500$ blockwise, vs hard projection 0.321–0.338), but the per-game statistics differ. Softplus chaotic-game median exponents are $2.7\times$ lower at 7.3×10^{-5}, with lower per-block sign stability (median positive-block fraction 0.83 vs 1.00, all-blocks-positive 0.27 vs 0.66). That measured 2.7-fold difference exceeds the two-fold bound stated before the runs. Per-game exponent statistics differ, and the slow chaos of the softplus reparameterization is more intermittent at long horizon. At each late-run reorthonormalization our time-local companion measurement records the interval growth of the leading Lyapunov vector together with a low-multiplier indicator ($\lambda < 0.5$ for either player) and accumulates per-game conditional means and lagged cross-correlations. Table A6 has the timing rows. We use 0.5 to label low-multiplier periods. At the peak cell the median active multiplier is 4.3. This puts 0.5 at the lowest $\sim 12\%$ of the median active multiplier (5.9% of active-multiplier mass lies below it). The threshold enters only the labeling. Tangent expansion is not concentrated in low-multiplier periods and is mildly suppressed just after a visit there. A non-binding expectation of a positive lagged peak, stated before the runs, is not borne out. Visits to the low-multiplier region set where trajectories go on time scales slower than the oscillation.

D.8.9 Frequency measurement

A censored-regression variant of the frequency fit treats oscillators with frequency within three times the resolution limit as left-censored (Gaussian maximum likelihood in log-log coordinates). It gives slope 0.897 with scatter 0.17 decades. The monotone sequence is 0.823 (all oscillators), 0.860 (censored oscillators dropped), 0.897 (censoring modeled). Figure 4 (main text) is the measured-vs-predicted scatter. Table A7 records the per-cell slopes with bootstrap 95% intervals, residual medians, and inclusion rates. Slopes are 0.99–1.09 in the $K = 2$, $\tau \geq 1$, $\zeta = 0.8$ cells. Intervals contain one at $\tau = 1$ for all three step sizes, while at $\tau = 10$ the interval $[1.04, 1.09]$ lies just above one. Slopes flatten in the small-step-size-ratio cells, the cells with slack slightly below zero, and the $K = 8$ cells. Residual medians are largest there too. An alternative predictor $\eta\sqrt{\tau}$ (variance term dropped) has pooled log-log $r^2 = 0.69$ (0.91 for the frequency formula). That predictor is constant within cells, while the measured frequency correlates with $\sqrt{\mathrm{Var}_x(c)}$ at median 0.90 within cells. Median log-residuals are $0.01/0.05/0.01$ across $\eta = 0.02/0.05/0.1$, $0.42/0.03/0.02$ across $\tau = 0.1/1/10$, and $0.04/0.04$ across $K = 2/8$.

D.8.10 Thresholds and resolution

The movement and rate thresholds are the same absolute values at every N and S. The variance normalization of the N-player construction keeps gradient scales matched across N (Appendix D.7), and the Markov variant sums the step norm over states as stated. The threshold

Table A7: Stratified frequency-formula statistics: per-cell log-log ordinary least squares (OLS) slope of measured on predicted frequency with bootstrap 95% intervals, median log-residual, and inclusion rate (included oscillators over $2G$). Cells with fewer than 30 included oscillators are omitted from the slope rows but counted in the pooled fit.

η	τ	K	ζ	n	slope	95% interval	med. resid.	incl.
0.02	1	2	0.8	73	0.99	[0.86, 1.17]	+0.01	0.036
0.05	0.1	2	0.8	73	0.60	[0.49, 0.79]	+0.12	0.036
0.05	0.1	2	1	323	0.56	[0.53, 0.60]	+0.46	0.162
0.05	0.1	8	0.8	122	0.59	[0.50, 0.69]	+0.39	0.061
0.05	1	2	0.8	142	1.07	[0.98, 1.17]	+0.01	0.071
0.05	1	2	1	305	0.80	[0.77, 0.83]	+0.22	0.152
0.05	1	8	0.8	180	0.76	[0.66, 0.88]	+0.04	0.090
0.05	10	2	0.6	34	1.09	[1.04, 1.17]	-0.06	0.017
0.05	10	2	0.8	282	1.07	[1.04, 1.09]	+0.00	0.141
0.05	10	2	1	544	0.84	[0.78, 0.89]	+0.05	0.272
0.05	10	8	0.8	548	0.86	[0.75, 0.95]	+0.02	0.274
0.1	1	2	0.8	164	1.04	[1.00, 1.08]	+0.01	0.082

Table A8: Classifier operating characteristics per budget-ratio cell against the $T = 3 \times 10^5$ tangent-map estimates ($\rho = 0$, $\tau = 1$).

ζ	0.5	0.6	0.7	0.75	0.8	0.9	1.0	1.2
recall	0.50	0.91	0.94	0.97	0.94	0.94	0.90	0.81
precision	0.03	0.18	0.35	0.41	0.40	0.43	0.40	0.38

audit of Table A33 bounds how far a tenfold change in either threshold moves the labels. Our growth threshold is calibrated on a known-answer family: skew-tent maps with exact top exponent $-a \ln a - (1-a)\ln(1-a)$, tuned to multiples of the threshold $3/(T/2)$ and run under the paper's exact accumulation protocol. Those maps are detected at rates $0.22/0.18$ ($T = 3 \times 10^4/3 \times 10^5$) when the true exponent equals the threshold and at $0.68/0.60$ at four times it, while a zero-exponent rotation is never labeled chaotic. Detection depends on the increment distribution: the rare large increments of the skew-tent map are the hardest case, and Gaussian-increment growth is detected at 50% at threshold by symmetry. A game counts as chaotic when its accumulated top-exponent log growth over the second half of the run exceeds 3. The horizon sequence of Appendix D.1 uses 1,000 games per cell. Its broad-chaos criterion fixed before the runs was a chaotic fraction reaching 0.40. It is not met, since the fraction saturates at 0.32–0.34.

Our frequency pipeline measures the spending $x^\top c$ per player over the final 18,000 steps. Here ω is the dominant peak of the Hann-windowed periodogram with log-parabolic sub-bin interpolation, taken after the mean is removed from the spending series. It is an angular frequency in radians per step, the convention of Lemma 1. The prediction uses $\mathrm{Var}_x(c)$ averaged over the same steps. The resolution limit marked in Figure 4 (dotted) is the minimum measurable frequency over those final steps.

D.8.11 Horizon checks and checks in other rows

The horizon extrapolation of the fine ζ sweep reruns every sweep cell at $T = 120{,}000$ on the sweep's seeds. Per-cell separating fractions move by at most 0.06 (converged fractions erode slowly, as the trapping audit predicts). The peak stays at $\zeta = 0.7$–0.75. Tangent-map estimates off the $\rho = 0$ row use $T = 300{,}000$ and paired classifier runs. At $\rho = -0.5$ the chaotic fraction is 0.13, 0.44, 0.48, 0.38 at $\zeta = 0.6, 0.75, 1.0, 1.2$, an interior peak near $\zeta = 1.0$. This lies outward of the classifier peak at 0.75, where the classifier fraction is 0.53, 0.84, 0.69, 0.48. At $\rho = +0.5$ the chaotic fraction is 0.07, 0.20, 0.12, with the peak at 0.75 as for the classifier. The chaotic-fraction peak exists in both rows. Table A8 gives the per-cell operating characteristics of the classifier against the 300,000-step tangent-map estimates: high recall at every intermediate budget ratio, precision near 0.4 there and collapsing in the tightest-budget cell where almost no game is chaotic.

D.8.12 Stability of the chaotic fraction

The chaotic fraction at the peak cell is stable to initialization, burn-in, and blocking. Table A6 collects the audits. Population fractions are stable across starts, and per-game labels mostly so. A

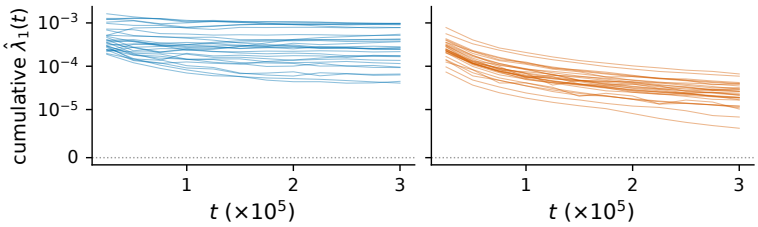


Figure A2: Cumulative top-exponent estimates $\hat{\Lambda}_1(t)$ per game (symlog scale), 30 games sampled per panel from the peak cell at $T = 3 \times 10^5$. Left: chaotic games. Right: games that are not chaotic.

Table A9: Event-aligned profile of visits to the low-multiplier region over chaotic games (per-interval growth of the leading Lyapunov vector; intervals of 20 steps; medians).

System	approach growth	post-entry growth	within-visit displacement	events/game
hard projection	0.019–0.023	0.0045–0.006	1.26	18
softplus reparameterization	0.013–0.015	0.005–0.008	1.51	9

boundary population of about 20% depends on the initialization. The chaotic fraction falls from 0.343 to 0.292 as burn-in increases, consistent with the exponents of the games that are not chaotic shrinking rather than crossing the growth threshold. Slow chaos is intermittent at block scale but stable in sign. Most chaotic games stay positive across blocks. Table A17 reports the decaying and sign-mixed cases. Figure A2 shows per-game cumulative estimates: chaotic games flatten onto positive plateaus, and games that are not chaotic decay toward zero.

The softplus reparameterization has chaotic fraction 0.352 against 0.321 for the primal-dual learner when re-measured at the same horizon and sample size as the peak cell ($T = 3 \times 10^5$, $G = 1{,}000$; the $G = 1{,}000$ counterpart of the 0.362 above). The chaotic fraction also persists under the step-size limit at matched $\eta^2 T$. At $\eta = 0.02$ with $\eta^2 T$ fixed ($T = 1.875 \times 10^6$, same games), the chaotic fraction is 0.327 vs 0.321 with per-game agreement 0.93. Exponents of chaotic games scale down by a median factor 0.38 for this pair. Across the fourfold step-size range of the collapse below, the chaotic-game median exponent scales between linearly and quadratically in η (factors 1.9–3.6 per doubling).

D.8.13 Expansion before contact with the low-multiplier region

The measurement before contact tracks per game the segment between leaving the low-multiplier region ($\lambda \geq 0.5$ for both players) and first returning to it. Tangent QR growth is accumulated as two totals, one inside and one after that segment ($\zeta = 0.75$, $G = 1{,}000$). Games that never leave the initial dwell (100) or never return (50) are excluded with counts. Among the 136 chaotic games with both segments, 99.3% accumulate net positive tangent growth already before the first return to the low-multiplier region. The median rate is 4.1×10^{-3} per step over a median 659-step segment, against the long-run chaotic-game exponents of 10^{-5}–10^{-3}. The squared-norm share of the leading Lyapunov vector in that segment is 0.998 in the strategy block. Expansion neither starts at nor waits for contact with the low-multiplier region. This matches the independent-phase average of Proposition 2. The low-multiplier region decides reinjection on longer time scales. An event-aligned measurement traces a visit to the low-multiplier region, with entries detected at renormalization states and offsets of 20 steps. Growth of the leading Lyapunov vector peaks on the approach, collapses three- to five-fold in the first intervals after entry, and recovers by ten intervals. Meanwhile low-probability-action identity switching stays flat (0.13–0.17) and the strategy state moves a large within-visit distance (Table A9). Contraction, folding, and large reinjection happen within a visit, while the energy growth happens on the approach. This is the within-visit counterpart of the measurement before contact. The softplus reparameterization has no kink at all. The same analysis there reproduces the profile with the low-multiplier region defined functionally as $\lambda = \mathrm{softplus}(z) < 0.5$. The low-multiplier region acts through low-multiplier saturation and the multiplier's integral of past violations and not through the kink. The softplus reparameterization locates the folding in that region.

D.8.14 STEP-SIZE DEPENDENCE

Decaying step sizes behave as the η^2 scaling of the independent-phase average predicts. The step is $\eta_t = 0.05/(1 + t/3000)^p$ at $T = 120{,}000$ with separation rates computed over thirds of the run (thresholds are not reused, since movement scales with the step). A positive separation rate in the final third persists for $p = 0.5$, where $\sum_t \eta_t^2$ diverges logarithmically: the positive-rate fraction is 0.53 at the peak cell vs 0.63 for constant steps. The rate stalls for $p = 1$, where the sum converges (0.26, and 0.09 at $\zeta = 0.9$). The separating-fraction peak needs constant or slowly decaying steps. That $p = 1$ stall is freezing rather than convergence: the final states of feasible games have median exploitability 0.21 (90th percentile 1.2; 23% below 10^{-2}) with small violation of median 0.004. Our step-size triple (full budget-ratio sweeps at $\eta = 0.02, 0.05, 0.1$ with $\eta^2 T$ fixed) shows the peak location stable at $\zeta = 0.75$ under all three step sizes. Outcome fractions themselves retain a step-size dependence that grows toward looser budgets: the maximum per-cell fraction deviation is 0.02 at $\zeta = 0.5$, 0.13 at the peak cell, and 0.36 at 1.2. The separating fraction increases with η.

D.8.15 SCALING COLLAPSE OF THE TANGENT-MAP ESTIMATE

For the tangent-map estimate the collapse is sharper ($\eta \in \{0.025, 0.05, 0.1\}$, drift scale $s = \eta^2 T \in \{187.5, 750\}$, $\zeta \in \{0.6, 0.75, 0.9, 1.2\}$, $G = 300$ per cell). At fixed s the chaotic fractions stay within ± 0.05 of their cell mean across the fourfold step-size range at every budget ratio. Their largest pairwise spread is 0.07. The peak cell gives 0.32/0.27/0.25 at $s = 187.5$ and 0.34/0.34/0.27 at $s = 750$ for $\eta = 0.025$/0.05/0.1. The peak position never moves, and the all-game median of the accumulated growth $\Lambda_1 T$ matches cell by cell (1.6–1.9 at the peak cell, 0.2–0.8 at the loosest budget). The median $\Lambda_1 T$ of the chaotic games falls with η (29.8, 22.4, 14.8 at $s = 187.5$). Across s the fractions rise with s alone, the known horizon dependence. The residual step-size dependence of the classifier above belongs to the classifier and not to the Lyapunov spectrum.

D.8.16 STATISTICAL CONVENTIONS

Cross-condition comparisons use matched game draws wherever seeds are shared (the paired-seed design of the intervention, update, and control variants). Matched-variant differences therefore have paired uncertainty tighter than the quoted per-cell intervals. Those are the conservative unpaired ones. Quoted intervals cover game-draw sampling; additional audits vary horizon, initialization, classification thresholds, and matched variants. The ± 0.034 per-cell interval is a conservative two-sided 3σ bound for one Bernoulli fraction at $G = 2{,}000$. Every confirmatory verdict rests on fresh-seed reruns. Under a global null the 3σ threshold gives 0.12 expected false positives across the 45 coarse-grid cells. At the peak cell, the separating fraction is 0.763 $[0.757, 0.769]$ against 0.027 $[0.025, 0.029]$ for the classifier in the no-budget cell. The chaotic fraction is 0.338 $[0.317, 0.359]$ against 0.044 $[0.036, 0.054]$ in the same no-budget cell with the same estimator and horizon (95% Wilson intervals). Differences between the peak cell and the no-budget cell are 0.74 and 0.29 with non-overlapping intervals, each contrast within one measurement. That no-budget comparison is also state-matched. A control with a loose budget ($\zeta = 5$, multipliers present but inert, $G = 500$, $T = 3 \times 10^5$) gives 0.056, the no-budget value 0.044 within noise. Experiments divide into confirmatory and exploratory. A confirmatory experiment had its criterion fixed before launch in a time-stamped log. That is what fixed before the runs means throughout this paper. The confirmatory set is: the robustness-set verdicts of Table A36 (whose criteria are the peak and tightest-cell definitions in its caption); the threshold, upper-bound, and simultaneous-update audits; the long-run ζ sweeps; and the checks in other rows. The exploratory set is reported descriptively: the mechanism-localization measurements (whose embedded sub-checks had criteria stated before the runs and are judged where reported), the stratified regressions, and this appendix's decompositions. The full set of unmet confirmatory predictions is: the τ-aggregate frequency check; the $\rho = -0.5$ peak-location comparison; the simultaneous-update expectation; the horizon sequence's broad-chaos bar of 0.40; the lead-lag majority-lead prediction; the cross-distribution peak at wide support (a cell-level miss); the reduced-return-map variant (area under the ROC curve (AUC) 0.59 vs the required 0.70); the expectation that the cost direction alone would lose positivity (unmet in the 75% that retain it); the softplus two-fold exponent bound (missed, recorded with its measurement); the always-feasible-budget peak comparison (misses its criterion, recorded as a decomposition); the expectation of a peak lagging visits to the low-multiplier region (non-binding, not borne out); and the sliding-window monotonicity expectation (holds through $W = 1000$ with a recorded non-monotone tail at $W = 5000$).

Table A10: Frequency-formula selection sensitivity: the pooled fit under no filtering, the stated inclusion rule, and additional censoring of oscillators within three times the resolution limit.

Rule	n	median ratio	pooled slope	r^2
no filtering (nonzero predictions)	3,291	1.047	0.59	n/a
stated inclusion rule	2,801	1.040	0.82	0.91
rule + censoring near the resolution limit	2,454	1.03	0.86	0.88

D.8.17 FREQUENCY INCLUSION RULE

An oscillator (one player in one game) enters the pooled regression when the game classifies as cycling; the mean multiplier lies in $(0.05,\ 0.95\,\lambda_{\max})$; $\mathrm{Var}_x(c) > 10^{-6}$; and the dominant peak holds at least 10% of the spectral power. Inclusion-composition bias in Section 5 is how the included population changes across τ cells. Across the sweep the mean class fractions are 0.55 converged, 0.19 cycling, and 0.27 separating. Only cycling games enter. Among recorded cycling oscillators 54% are included. The multiplier range drives most exclusions: 94% lie outside the oscillating working range, mostly with multipliers at zero; 81% overlap with near-zero cost variance; weak spectral peaks account for 1%. The median $|\log_{10}|$ error of the included oscillators is 0.029. Inclusion rate by predicted-frequency quartile is $0.00/0.19/0.86/0.91$, because low predicted frequencies fall below the resolution limit of the series. The $2{,}345$ included oscillators whose predicted frequency lies in the fixed range every series can resolve ($5\times$ the resolution limit to a tenth of Nyquist) give slope 0.92, correlation 0.92, and median measured-to-predicted ratio 1.02. Our spectral-peak estimator is calibrated on synthetic known-frequency sinusoids with matched amplitude and noise: median absolute log error 0.007, 95th percentile 0.11. The stratified flattening therefore reflects resolvability rather than estimator bias.

The fixed-point conditions of the two-action theory hold in 24% of the two-action draws. Over the $K = 2$ ensemble at $\zeta = 1$, the interior strategy that meets the budget with equality exists for both players in all games (a property of the uniform-play-cost budget at $K = 2$). Those conditions are that both fixed-point multipliers are nonnegative at that profile, against a measured cycling fraction of 0.28. We chart the population of the independent-phase average the same way. Competitive coupling ($u'v' < 0$) holds in 0.49 of $K = 2$ draws, with separating fractions 0.43 competitive against 0.36 coordination. On the subset where the fixed point exists and the coupling is competitive, the separating fraction is 0.75. The e-fold time of the energy growth has median $\sim 3{,}100$ steps against a ~ 250-step oscillation period. The energy growth accumulates over roughly a dozen oscillations: slow against the period and fast against the horizon.

The rule retains $2{,}801$ of $5{,}186$ measured cycling oscillators across the frequency sweep. The cells of the sweep were chosen before the fit. Per-cell inclusion fractions lie between 0.15 and 0.66 of the cycling population, lowest in the tightest-budget cell with the small step-size ratio. Per-cell counts are in the supplementary records.

We also predict the late-segment frequency from the variance over an earlier segment and from the uniform-strategy cost variance. The sweep is rerun with a split record and $T = 60{,}000$. The variance measured on an early segment (steps $12{,}000$ to $30{,}000$) predicts the late-segment frequency at within-cell correlation 0.63, the median over cells. A fully ex ante predictor that uses only the cost vector (variance under the uniform strategy, no trajectory information at all) reaches 0.75 with a stable scale offset. The median measured-to-predicted ratio is 0.78 because the equilibrium strategy is less spread than uniform. Cross-game frequency ordering is thus largely predictable from the game instance alone. Tail variance gives the highest within-cell correlation, 0.90. By construction the rule excludes converged games (nothing oscillates) and separating games with no stable frequency. These are 73–99.8% of oscillators depending on the outcome mix of the cell. Within the cycling population, inclusion is 0.43–0.66 in every cell except the tightest-budget one with the small step-size ratio (0.15), where oscillations lie at the resolution limit. We re-fit under three rules to measure selection sensitivity (Table A10). Unfiltered and filtered medians nearly coincide, with 80% of unfiltered oscillators within a factor of two. The pooled slope degrades without filtering as oscillators at the resolution limit and oscillators with spurious peaks enter, and improves when oscillators near the resolution limit are censored.

D.8.18 EXPLOITABILITY

The feasible-deviation component of the exploitability in Appendix D.4 evaluates each player's feasible best response for a profile (x, y), with $q = Ay$ for player 1 and $q = B^\top x$ for player 2: $\max\{q^\top x' : x' \in \mathcal{S}_K,\ c^\top x' \le d\}$. It does so exactly by vertex enumeration, since the optimum has support at most two (feasible single actions and two-action mixes that meet the budget with equality). This is verified against an LP solver to 10^{-15}. The component reports the sum over players of the positive parts of the best-response payoff minus the current payoff. The positive part matters because trajectory strategies are not confined to the budget set. A strategy with positive violation can earn more than the feasible best response, making the raw difference negative. The fraction of games where this occurs is reported as a diagnostic. Raw differences are in the per-game records. Both the last iterate and the tail-averaged profile (both players averaged over the final 20%) are evaluated. Exploitability is defined for feasible games alone, and infeasible games are excluded with counts reported.

D.9 ROBUSTNESS AUDITS OF THE PROTOCOL

The interior peak persists across the tested classifier thresholds, upper bounds, and the two updates. What changes relative to the main protocol is one protocol choice or one construction at a time, on the same drawn games and seeds wherever the variant allows it.

D.9.1 VISITS TO ZERO AND THE SEPARATING LABEL

Visits to zero alone do not produce the separating label: periodic orbits classify as cycling. Near-universal visits to zero coexist with separating fractions from 0.058 to 0.977 across the constrained cells of the coarse grid (Section 4.2; Table A34). Tight-budget cells with over half the multipliers at the upper bound are re-run at ten times the upper bound, and the upper-bound scan below confirms invariance to the upper bound at and above the default.

D.9.2 CLASSIFIER THRESHOLD AUDIT

Table A33 reclassifies every cell of the fine ζ sweep (and its no-budget cell) from the frozen per-game records under all nine combinations of the convergence threshold $\{10^{-5}, 10^{-6}, 10^{-7}\}$ and the separation-rate threshold $\{10^{-2}, 10^{-3}, 10^{-4}\}$ (Section 5). The peak persists in every combination. The peak is at $\zeta = 0.70, 0.75, 0.80$ under rate thresholds 10^{-2}, 10^{-3}, 10^{-4} respectively (heights 0.48, 0.76, 0.88; the convergence threshold changes nothing). This one-cell shift tracks classifier sensitivity while the interior position of the peak is stable across thresholds. The loosest rate threshold lifts the no-budget cell from $\approx 3\%$ to 47% separating (the matched chaotic fraction there is 0.044).

D.9.3 THRESHOLD-FREE PRESENTATION

The continuous separation rate of the classifier has per-cell medians 0.0004, 0.010, 0.013, 0.010, 0.005, 0.001 and 90th percentiles 0.019, 0.051, 0.069, 0.068, 0.053, 0.031 at $\zeta = 0.5, 0.6, 0.7, 0.75, 0.85, 1.0$ (20,000 games per cell). The top exponent of the tangent-map estimate has medians 0, 0, 5.2×10^{-6}, 3.4×10^{-6}, 9.0×10^{-7}, 5.6×10^{-7} with 90th percentiles 0, 1.2×10^{-6}, 2.8×10^{-4}, 4.1×10^{-4}, 2.5×10^{-5}, 1.0×10^{-5} at $\zeta = 0.25, 0.5, 0.75, 1.0, 1.25, 1.5$ ($T = 3 \times 10^5$). Both continuous distributions have an interior peak.

D.9.4 MULTIPLIER UPPER-BOUND SCAN

Table A11 covers the peak cell and its neighbors $\zeta = 0.7$ and 0.8: fresh-seed 20,000-game cells at $\lambda_{\max} \in \{12.5, 25, 100, 200, 500\}$ against the $\lambda_{\max} = 50$ cells behind Figure 3. The upper bound enters the dynamics only when it lies inside the oscillation range of the multiplier. From the default upper bound upward the fractions are flat and the fraction of games at the upper bound vanishes.

At $\zeta = 0.5$ most games are infeasible. The same scan run to $\lambda_{\max} = 500$ leaves the converged fraction unchanged cell by cell at 0.89 to 0.92, while the share of runs at the upper bound falls from 0.98 to 0.03 at $\tau = 1$. The convergence of infeasible games onto the cheapest action is therefore a

property of the dynamics and not of the bound. At $\tau = 10$ the share at the bound stays near 0.97 even at $\lambda_{\max} = 500$. This share is the multiplier step overshooting rather than the bound acting.

Table A11: Upper-bound scan near the peak cell at $\rho = 0$, $\tau = 1$, 20,000 games per cell: classifier separating fraction with the fraction of games at the upper bound in parentheses. The $\lambda_{\max}$=50 column is the baseline from the fine ζ sweep.

$\zeta \setminus \lambda_{\max}$	12.5	25	50	100	200	500
0.7	0.69 (.37)	0.72 (.27)	0.74 (.19)	0.76 (.10)	0.75 (.02)	0.76 (.00)
0.75	0.73 (.19)	0.75 (.12)	0.76 (.08)	0.77 (.03)	0.77 (.00)	0.77 (.00)
0.8	0.72 (.09)	0.74 (.04)	0.74 (.02)	0.75 (.01)	0.75 (.00)	0.74 (.00)

D.9.5 HOLD-AT-ZERO INTERVENTION

The hold-at-zero variant holds a player's multiplier at zero for R further steps on each visit to zero (positive multiplier projected to zero) and leaves the dynamics everywhere else unchanged. This lengthens residence in the low-multiplier region without changing local gains or the multiplier's integral of past violations outside the hold. $R = 0$ is bitwise identical to the primal-dual learner. Post-warm residence at zero rises from 0.278 at $R = 0$ to 0.287, 0.316, 0.363 at $R = 100$, 300, 1000 (peak cell, $G = 2{,}000$). The separating fraction stays flat through $R = 300$ at 0.763, 0.768, 0.770 and falls only at $R = 1000$ (0.688, with convergence and cycling both rising). That hold outlasts the ≈ 600-step oscillation period. Residence in the low-multiplier region is therefore not itself the sensitive quantity. The raised-lower-bound scan shortens the visits to zero without touching gains. The region therefore matters through its recurrence geometry and the multiplier's integral of past violations and not through the delay it imposes.

D.9.6 INTERVENTIONS AT ZERO

Table A12 reports the interventions at zero of Section 5, with all variants sharing drawn games and seeds ($G = 2{,}000$ per cell). The softplus-smoothed projection bottoms out at a height of order $1/\beta$, while the softplus reparameterization has no non-smooth point and lets the multiplier approach zero.

Table A12: Interventions at zero for $\rho = 0$, $\tau = 1$, 2,000 games per cell, matched seeds: separating/cycling fractions. At $\zeta = 0.75$ the softplus-smoothed projection reads $0.31/0.16$ at β=2 and $0.73/0.12$ at β=32.

ζ	0.6	0.7	0.75	0.8	0.9	1.0
hard projection	0.48/0.05	0.73/0.09	0.77/0.11	0.72/0.16	0.65/0.23	0.54/0.29
reparameterized	0.46/0.01	0.74/0.02	0.79/0.02	0.79/0.04	0.72/0.09	0.60/0.18
smoothed (β=8)	0.41/0.06	0.63/0.11	0.64/0.13	0.59/0.18	0.48/0.22	0.35/0.28

Table A13 collects the record of interventions at zero and on the multiplier's integral of past violations. The variant with the C^1 lower-bound map tracks the hard projection on the same drawn games. The raised-lower-bound sweep leaves the peak nearly unchanged through $\lambda_{\min} = 0.3$ and erodes it only at 1. The exact bound at zero is immaterial, and any region of small multipliers below the median active multiplier suffices. The penalty variant removes the peak entirely at every gain and budget ratio: a bound at zero without the multiplier's integral of past violations does not produce the peak in the separating fraction. The sliding-window multiplier makes the length of the integral of past violations continuous. It applies the matched ramp gain $\tau\eta$ over the last W violations with projection of the output, and $W = 1$ coincides bitwise with the penalty variant. The peak weakens for window lengths shorter than the oscillation period ($2\pi/\omega \approx 600$ steps at the peak cell), strengthens as the window length crosses it, and reaches 78–94% of the fraction for the primal-dual learner by $W = 1000$. Our monotone expectation stated before the runs holds through $W = 1000$. A recorded non-monotone value appears at $W = 5000$, where the output-projected sliding-window multiplier departs from the state-projected integral (construction in Appendix D.7).

The multiplier's integral of past violations therefore has a scale, the oscillation period, below which the peak in the separating fraction cannot form. The sliding-window multiplier holds the ramp gain fixed. That isolates the window length from gain.

Table A13: Interventions at zero and on the multiplier's integral of past violations, classifier separating fractions on matched drawn games ($G = 2{,}000$ per cell). The first two rows are at $\zeta = 0.7/0.75/0.8$ and the $\lambda_{\min} = 1$ row at $\zeta = 0.6/0.75/0.9$; the remaining rows are at the cells their labels name.

Variant	Separating fractions
Hard projection (baseline, cells of the C^1 lower-bound map)	0.73 / 0.77 / 0.72
C^1 lower-bound map, $\delta = 0.05$ and 0.2	0.73 / 0.75 / 0.71 and 0.70 / 0.72 / 0.69
Lower bound $\lambda_{\min} = 0$ to 0.3 (peak cell)	0.768, 0.764, 0.758, 0.745, 0.713 (same flatness at $\zeta = 0.6, 0.9$)
Lower bound $\lambda_{\min} = 1$	0.403 / 0.606 / 0.429
Penalty variant (all gains, all ζ)	$\leq$ 0.023 everywhere, against 0.77 for the unchanged learner
Sliding-window multiplier $W =$ 1/50/200/1000/5000, $\zeta = 0.75$	0.03 / 0.02 / 0.33 / 0.60 / 0.54 (unchanged learner 0.77)
same at $\zeta = 0.6$ and 0.9	0.04/0.01/0.26/0.46/0.36 (unchanged learner 0.49) and 0.02/0.02/0.32/0.60/0.52 (unchanged learner 0.65)

D.9.7 Exploitability record

Table A14 gives the per-cell medians of the feasible-deviation component. The converged class is heterogeneous: the share of converged games meeting the success criterion (small exploitability and small violation) falls from 0.90 at looser budgets to 0.44 near the peak. Across the three cells near the peak ($\zeta = 0.7$–0.8) the feasible separating class totals 4,050 games, of which 6 jointly reach a small exploitability and a small violation. Among converged games the fraction with exploitability and violation jointly below 10^{-3} is 0.31, 0.44, 0.52, 0.68, 0.78, 0.84, 0.90, 0.90 at $\zeta = 0.6$ through ∞ in the order of the table. The equilibrium subclass dominates everywhere except near the tightest budgets. There upper-bound-frozen and penalty-frozen rest points and slow near-rest trajectories carry exploitability up to order one under the convergence label.

Table A14: Exploitability (feasible-deviation component) at $\rho = 0$, $\tau = 1$, 2,000 games per cell, feasibility-conditioned medians: last-iterate / tail-averaged exploitability per classifier label, and tail violation of the separating class. At $\zeta = 0.5$ only 6 feasible separating games exist (0.3% feasible), so that cell is omitted.

ζ	0.6	0.7	0.75	0.8	0.9	1.0	1.2	∞
converged	.018/.018	0/0	0/0	0/0	0/0	0/0	0/0	0/0
cycling	.015/.0001	.045/.0005	.043/.0009	.043/.0007	.100/.0015	.117/.0026	.266/.0079	.795/.039
separating	.189/.018	.265/.024	.304/.024	.318/.024	.361/.022	.411/.027	.472/.028	1.62/.192
tail violation	.024	.031	.035	.036	.036	.036	.020	0

Our averaging-horizon scan behind Appendix D.4 evaluates the exploitability of the full-prefix averaged profile at horizons 3,750 to 120,000 on the same drawn games and classifier labels. Pooled over the cells at $\zeta = 0.75$ and 0.8, the cycling-class median falls 0.039, 0.021, 0.011, 0.006, 0.003, 0.002 across the doubling horizons while the separating-class median falls 0.046, 0.036, 0.028, 0.022, 0.016, 0.012. In the no-budget cell the separating class is flat at 0.026 from the third horizon on.

D.9.8 Classifier audits and sampled feedback

The renormalized coordinates and the renormalization interval of the classifier are audited directly. Renormalizing the full state (logits and multipliers) instead of the strategy logits preserves the interior peak, with fractions 0.001–0.15 peaked at $\zeta = 0.9$–1.0. Absolute rates are lower because the contracting multiplier directions dilute the normalized displacement. Renormalization intervals 25 and 100 reproduce the fine ζ sweep (0.78 and 0.72 at the peak cell vs 0.76). Perturbation direction is immaterial: three alternative directions move the separating fraction at the peak cell by at most 0.005, from 0.761 to 0.770, on the same drawn games. Table A15 collects the sampled-feedback measurements, including the tail movement under one-sample feedback. Under one-sample feedback with common random numbers the separating fraction is zero at every budget ratio while the tail movement stays elevated around the peak. Sampling only the payoff or only the cost of the multiplier does not remove the separating fraction, because the feedback left exact keeps the state dependence of its block of the Jacobian. Both the classifier and the tangent-map estimate through the

Table A15: Sampled-feedback record ($\rho = 0$, $\tau = 1$; separating fractions under common random numbers unless stated). The tail-movement row records the non-convergence that persists under sampling.

Measurement	Values
Separating fraction under common random numbers	0.000 at every budget ratio
Median tail movement, $\zeta = 0.5$ to 1.2 and ∞	0, 2.4×10^{-4}, 1.5×10^{-3}, 2.2×10^{-3}, 2.6×10^{-3}, 3.5×10^{-3}, 3.3×10^{-3}, 1.9×10^{-3}, 7.6×10^{-16} (fraction at the upper bound 0.98, 0.66 at the two tightest-budget cells)
Batch averaging at $\eta = 0.05$, $b = 1$ to 4096 ($\zeta = 0.75$)	0.000, 0.000, 0.000, 0.001, 0.003, 0.004, 0.021 (0.022 at $\zeta = 0.8$, $b = 4096$); tail movement flat in b
Batch onset at other step sizes	$\eta = 0.1$: 0.003, 0.005, 0.024 at $b = 256$, 1024, 4096; $\eta = 0.02$: 0.013 already at $b = 1024$
Payoff or cost feedback sampled, the other exact ($\zeta = 0.75$ / 0.6 / 1.0, $b = 1$): payoff only; dual cost only	0.428 / 0.478 / 0.110; 0.299 / 0.044 / 0.295 (exact feedback 0.748 / 0.483 / 0.532; both sampled 0.000); tail movement unchanged; at $b = 1024$, $\zeta = 0.75$: 0.411; 0.379 (both sampled 0.0035)
Tangent-map estimate through the sampled map ($G = 200$, $T = 3\times10^5$), chaotic fraction	both sampled 0.000 at every budget ratio; payoff only 0.450 / 0.110 / 0.165; dual cost only 0.505 / 0.130 / 0.515

sampled map stay positive at a level comparable to exact feedback, whatever the batch size. Under batch averaging the separating fraction re-emerges only faintly, with an onset non-monotone in the step size. Copies driven by independent noise streams test positive on the separation test in every game, as the construction implies.

D.9.9 The pathwise cocycle exponent is structurally exactly zero

Along any realized noise path the sampled indices enter through inverse-CDF comparisons and are locally constant in the state. The almost-everywhere Jacobian of that map is therefore triangular in (logit, multiplier) coordinates. The strategy-logit block is the identity because the sampled gradient is locally state-independent, and the multiplier diagonal is the indicator that the projection leaves the update unchanged. For triangular cocycles the Lyapunov exponents are the Birkhoff averages (time averages along the trajectory) of the log-diagonals. The top pathwise exponent is therefore exactly 0 for any finite batch under this inverse-CDF common-random-number realization, with negative values only through steps where the projection changes the update. A tangent-map estimate through the realized one-sample map under common random numbers confirms it. The trajectory reproduces the primal-dual learner under sampled feedback, and the linearization matches finite differences under the same common random numbers at non-flipping steps. Top-exponent median is 0.0 and the 95th percentile is $\leq 5\times10^{-6}$ at the peak cell, a looser-budget cell, and the no-budget cell ($G = 200$, $T = 3\times10^5$, $w = 3$ tangent vectors).

Table A16 records what sampling does to the outcomes on the same games. At the peak cell sampling both payoff and cost feedback moves the mean violation by $+0.0005$ and the median tail-averaged exploitability by $+0.005$ relative to exact feedback, and the other cells stay within 0.002 and 0.006. Sampling thus leaves budget satisfaction and equilibrium quality where exact feedback leaves them (criterion of 0.01 stated before the runs).

D.9.10 Averaging horizons and the four-way split

The feasible-deviation regret of the empirical joint distribution falls across the six horizons from 0.0093 to 0.0003 (cycling) and from 0.0004 to 0.0002 (separating) at $\zeta = 0.75$, from $0.0100/0.0041$ to $0.0004/0.0002$ at 0.8, and from $0.0229/0.0241$ to $0.0015/0.0017$ with no budget. In the averaging-horizon scan, reclassifying at $T = 120{,}000$ migrates 11.6% and 13.2% of the games in the two cells at $\zeta = 0.75$ and 0.8. Cycling and separating decay exponents are $0.91/0.39$ under the short-horizon labels, $0.90/0.37$ under the long-horizon labels, and $0.94/0.37$ on the label-stable subset. Joining tangent-map estimates at $T = 300{,}000$ to the same cells breaks these numbers down by the four-way split (Table A17). The last-iterate exploitability is monotone along the four-way split. The pooled separating exponent (~ 0.4) averages chaotic and not chaotic games, and time-averaging barely repairs the chaotic games.

Table A16: Outcomes under sampled feedback ($\rho = 0$, $\tau = 1$, $b = 1$, $G = 2{,}000$, $T = 30{,}000$, same games and initial logits per cell; feasible games): converged fraction, success rate (converged, tail violation at most 0.005, last-iterate exploitability at most 0.02), mean and 99th-percentile tail violation, median last-iterate and tail-averaged exploitability.

ζ	variant	converged	success	tail violation		exploitability	
				mean	p99	last iterate	tail-averaged
0.6	exact feedback	0.124	0.046	0.0225	0.072	0.147	0.0084
0.6	payoff sampled	0.132	0.049	0.0224	0.065	0.128	0.0141
0.6	cost sampled	0.144	0.054	0.0223	0.070	0.154	0.0113
0.6	both sampled	0.137	0.051	0.0225	0.068	0.150	0.0114
0.75	exact feedback	0.098	0.069	0.0311	0.095	0.195	0.0104
0.75	payoff sampled	0.097	0.068	0.0318	0.099	0.203	0.0139
0.75	cost sampled	0.098	0.067	0.0309	0.095	0.184	0.0142
0.75	both sampled	0.094	0.068	0.0317	0.099	0.220	0.0154
1	exact feedback	0.168	0.148	0.0251	0.100	0.138	0.0057
1	payoff sampled	0.166	0.145	0.0263	0.099	0.157	0.0110
1	cost sampled	0.169	0.150	0.0238	0.096	0.111	0.0072
1	both sampled	0.169	0.149	0.0253	0.097	0.147	0.0119

Table A17: Exploitability statistics by the four-way split at $\zeta = 0.75$ / 0.8 (labels of the tangent-map estimate joined per game): median last-iterate exploitability (feasible-deviation component) and averaged-marginals decay exponent.

Class	exploitability (0.75)	exploitability (0.8)	decay exponent
separating and chaotic	0.45	0.49	0.16 / 0.11
separating and not chaotic	0.21	0.20	0.67 / 0.64
cycling	0.043	0.043	0.93
converged	0	0	n/a

The chaotic fraction is 0.25, 0.29, 0.31, 0.32, 0.33, 0.33 at $T = 6 \times 10^4$, 1.2×10^5, 2×10^5, 3×10^5, 5×10^5, 10^6 on the same seeds and estimator (membership change $\pm 3\%$ between the last two, net -0.002). It is then 0.322 at 3×10^6 on the same games, with per-game label agreement 0.95 with 10^6 and a stable median exponent of the chaotic games (1.5×10^{-4} vs 1.8×10^{-4}). At 10^7 it is 0.34 on an independent 250-game draw. The marginally positive top exponents of games that are not chaotic shrink with horizon: median 3.2×10^{-6} at 3×10^5, 9.4×10^{-7} at 10^6, 2.6×10^{-7} at 3×10^6.

D.9.11 Three rules for the chaotic label side by side

The horizon sequence is re-read under three rules on the same per-game Lyapunov spectra, because the per-step growth threshold scales as $6/T$. Our growth threshold gives 0.25, 0.31, 0.32, 0.33, 0.32, 0.34 at six horizons from $T = 6 \times 10^4$ to 10^7. A fixed per-step threshold of 2×10^{-5}, the $T = 3 \times 10^5$ value, gives 0.65, 0.35, 0.32, 0.29, 0.27, 0.25. Its third entry coincides with the growth-threshold entry by construction. Short-horizon estimates are variance-inflated. A fixed absolute threshold shows the slow-chaos signature: the finite-time exponent distribution shifts slowly downward. Block certification gives 0.32 at $T = 3 \times 10^5$. The main fraction is stated at $T = 3 \times 10^5$ under the growth threshold, where all three rules give 0.32. Per-game label transition matrices across horizons are stable. From 3×10^5 to 10^6, 291 of 321 chaotic games stay chaotic (30 drop out, 37 enter). From 10^6 to 3×10^6, 301 of 328 stay (27 out, 21 in) on the same seeds. The two-way churn, about 3% of all games in each direction, is concentrated at the threshold boundary.

D.9.12 Intervention Lyapunov spectra

The interventions at zero and on the multiplier's integral of past violations are re-measured with the tangent-map estimate on the games of the peak cell ($G = 500$, $T = 3 \times 10^5$, seed-matched). In the penalty variant the multiplier is a function of state and is differentiated through by the tangent map. Chaotic fractions are collected in Table A18. Our tangent-map estimate reproduces the intervention ordering of the classifier exactly: the positive exponents themselves need the multiplier's integral of past violations and a reachable low-multiplier region and do not need the bound at zero.

D.9.13 MATCHED INTERVENTION DIAGNOSTICS

Same-game diagnostics across variants ($G = 300$, $T = 3 \times 10^4$, stride-10 series) distinguish an account through the multiplier's integral of past violations from a feedback-gain account. Post-warm mean multipliers are 5.0 (baseline), 1.4 (penalty variant), 0.5 ($W = 50$ window), 2.1 ($W = 1000$ window), 5.1 (raised lower bound 0.3), 4.9 (softplus). The fraction of games with an active multiplier is 0.49–0.78 across the free variants (the variant with a raised lower bound reaches 1.0 by construction), and mean violations run -0.04 to $+0.19$. A lower gain is thus shared alike by variants that remove the peak and variants that keep it. Within the sliding-window multipliers (one construction, one gain scale) the peak switches exactly as the window length crosses the oscillation period. The variants that remove the peak differ in the oscillation itself. The fraction of games with a dominant violation-frequency peak is 0.20 (penalty variant) and 0.26 ($W = 50$), against 0.51–0.56 for baseline, raised lower bound, softplus, and $W = 1000$.

D.9.14 VARIANT LYAPUNOV SPECTRA AND DECOMPOSITIONS BY BUDGET POSITION

Table A18 collects the Lyapunov spectra (tangent-map estimate) of the representative robustness variants and interventions: interior peaks in every variant. Small damping removes a third of the chaotic fraction at unchanged 99th-percentile tail violation before the tradeoff turns monotone.

The always-feasible budgets set $d = c_{\min} + s\,(\bar{c} - c_{\min})$ with $c_{\min} = \min_k c_k$, so s is a dimensionless budget position above the cheapest action: $s = 0$ affords that action alone and $s = 1$ affords uniform play. The cheapest vertex is affordable for every game at every s. On a 40-point grid ($s = 0.05$ to 2 in steps of 0.05, $G = 2{,}000$ per cell) the separating fraction declines from 0.84 at $s = 0.1$–0.15 to 0.09 at $s = 2$, with no interior local maximum above the 3σ interval anywhere on the grid (Figure A3). The single exception is the smallest budget position: $s = 0.05$ gives 0.72, below its neighbor, because the feasible set collapses toward the cheapest vertex and trapping at a pure strategy returns (convergence 0.26 vs 0.13 at $s = 0.1$). That drop at $s = 0.05$ is the feasible-geometry mechanism inside the always-feasible budgets, and the interval below $s = 0.1$ is the only one on the grid where the fraction increases with s. A paired re-measurement holds the drawn instances fixed across budget positions with one seed and budgets the only difference. It adds the tangent-map estimate ($T = 3{\times}10^5$, 500 games, nine budget positions; Table A19). Among instances with always-feasible budgets, stronger binding raises the separating fraction and the chaotic fraction monotonically under the classifier and the tangent-map estimate. The tight-budget part of the ζ axis belongs to feasibility geometry (removing it gives the tightest budgets the highest fractions). The decline at looser budgets belongs to the slack m, and the peak is their crossover. The paired tangent-map estimate gives 0.376 at $s = 0.1$, about $8.5\times$ the matched no-budget value 0.044. It declines to 0.054 at $s = 2$, statistically at the no-budget value. Binding is stronger at the smallest budget position of the always-feasible budgets than at the peak cell. At $s = 0.1$ the budget position $(d - c_{\min})/(\bar{c} - c_{\min})$ is 0.1 by construction. It is not the slack m. The median active multiplier there is 7.2 and the fraction of games with an active multiplier 0.96. At the peak cell the median budget position is 0.37, the median active multiplier 4.3, and the fraction with an active multiplier 0.91. These numbers agree with the feasibility-split reading, in which the peak of the fraction over

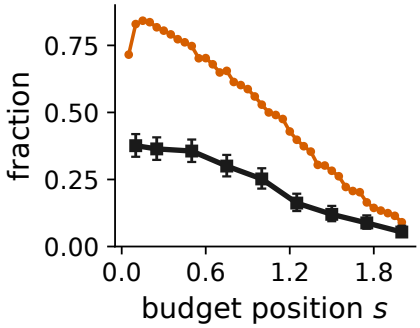


Figure A3: Always-feasible budgets $d = c_{\min} + s\,(\bar{c} - c_{\min})$ against the budget position s ($\rho = 0$, $\tau = 1$). Here s is 0 at the cheapest action and 1 at uniform play, not the slack m. Circles: separating. Squares: chaotic. Both fractions decline in s from $s = 0.15$ to $s = 2$, with no interior local maximum above the 3σ interval.

Table A18: Lyapunov spectra (tangent-map estimate) of robustness variants and interventions: chaotic fractions at the listed cells. Intervention variants share draws and seeds with the matched primal-dual learner. The sweep, two-budget, player, and Markov rows are fresh-seed fraction-level reads.

Variant	T	Chaotic fractions
Joint-action costs	1.2×10^5	0.09, 0.28, 0.15 at $\zeta = 0.6, 0.8, 1.2$
Plain-ascent reference (horizon-matched; the variants are matched in $\eta^2 T$)	swept	peak-cell 0.33 ($5{\times}10^5$ horizon), 0.29 ($1.2{\times}10^5$ horizon), 0.176 ($T{=}3{\times}10^4$, $G = 1{,}000$)
Extragradient, $\eta = 0.02/0.05/0.1/0.2$ at $\eta^2 T = 300$ ($G = 500/2{,}000/2{,}000/2{,}000$)	swept	0.06/0.24/0.09, 0.06/0.20/0.07, 0.04/0.16/0.06, 0.02/0.09/0.04 at $\zeta = 0.6/0.75/1.2$: the peak thins monotonically toward large steps but stays interior at every step size
Optimistic, same sweep and G	swept	0.06/0.21/0.09, 0.05/0.19/0.08, 0.03/0.15/0.05, 0.01/0.07/0.04: same pattern
Players $N{=}3$ ($K{=}8$) / $N{=}5$ ($K{=}4$)	3×10^5	peak-cell 0.64 / 0.65 vs tightest-budget 0.007 / 0.017 and no-budget 0.043 / 0.063 (tangent-map estimate, $G = 300$)
Markov variant ($S{=}4$, $K{=}4$)	3×10^5	0.005 / 0.605 / 0.33 / 0.29 at $\zeta = 0.5$ / 0.8 / 1.2 / ∞ (tangent-map estimate, $G = 200$: interior peak in the chaotic fraction, no-budget cell included)
$K = 32$	1.2×10^5	0.01, 0.49, 0.21 at $\zeta = 0.5, 0.7, 1.2$
Damped, $\delta = 0, 0.003, 0.01, 0.03$	1.2×10^5	0.29, 0.20, 0.15, 0.09 (p99 violations 0.10, 0.10, 0.11, 0.16)
Softplus reparameterization (ζ sweep)	1.2×10^5	0.064, 0.288, 0.297, 0.127 at $\zeta = 0.6$, 0.75, 0.9, 1.2
Softplus reparameterization (main protocol)	3×10^5	0.352 vs hard projection 0.321
Penalty variant	3×10^5	0.006
Payoff t_3 / uniform / sparse / rank-2	3×10^5	peak 0.280 / 0.293 / 0.185 / 0.250 vs own no-budget cell 0.040 / 0.048 / 0.003 / 0.013
Two independent budgets ($M{=}2$)	3×10^5	0.448 / 0.630 at $\zeta = 0.75$ / 0.9 (one-budget reference 0.32 / 0.31: the amplification seen by the classifier reproduces in the tangent-map estimate; expectation stated before the runs, met)
Duplicated budget ($M{=}2$)	3×10^5	0.335 / 0.278 at $\zeta = 0.75$ / 0.9 (doubled-gain control, at reference level as constructed)
Lower bound $\lambda_{\min} = 0.3$ / 1.0	3×10^5	0.286 / 0.172 (baseline 0.31)

all games lies at the feasibility transition and not at maximal binding. The chaotic fraction at the smallest budget position also exceeds the pooled fraction at the peak cell (0.376 vs 0.319, $z \approx 2.5$).

D.9.15 Per-game transitions across budget positions

The population decline is monotone but per-game membership is not. Across the nine same-seed budget positions 25% of the 396 games that are ever chaotic follow a strictly monotone chaotic-to-not-chaotic path, and adjacent budget positions exchange large in- and out-flows of chaotic games. Adjacent-position rank correlations of the top exponent run 0.32–0.77, weakest at the smallest budget positions. The identity of the chaotic games depends on the budget even where their number varies smoothly (weak static predictability, below).

D.9.16 Multiplicative always-feasible budgets

A second always-feasible construction with different geometry uses multiplicative budgets $d = \min_k c_k(1 + s)$ (20-cell grid, $G = 2{,}000$, monotonicity stated as the expectation before the runs). The separating fraction declines monotonically from 0.845 at $s = 0.15$ to 0.026 at $s = 1.95$ with

Table A19: Always-feasible budgets, same instances across budget positions s: paired-instance tangent-map estimate and classifier ($T = 3 \times 10^5$ / 3×10^4), and the classifier on the fresh-seed 40-point grid at the matching budget positions.

s	0.1	0.25	0.5	0.75	1.0	1.25	1.5	1.75	2.0
chaotic	0.376	0.364	0.356	0.300	0.252	0.162	0.120	0.088	0.054
paired classifier	0.815	n/a	0.749	n/a	0.522	n/a	0.286	n/a	0.090
fresh-grid classifier	0.830	0.818	0.748	0.655	0.529	0.398	0.283	0.165	0.091

no interior local maximum. The one exception to that decline is again the smallest budget position, where the fraction drops to 0.796 at $s = 0.05$, the same feasible-set collapse toward a pure strategy as in the always-feasible budgets above.

D.9.17 Curves conditioned on budget position across rows

Feasible games are binned by their per-game budget position (exploratory, frozen per-game records). The conditional separating fraction declines monotonically in budget position within every payoff-correlation row and step-size-ratio row, but the curves do not collapse onto a single function. At a fixed budget-position bin the level spans up to 0.5 across rows, ordered by payoff correlation (zero-sum highest, common-payoff lowest) and by step-size ratio. A per-game logistic regression of the separating label over the feasible games of the fine ζ sweep ($n = 247{,}035$) confirms the independent effect of budget position after adjusting for covariates. The standardized budget-position coefficient is -0.55: smaller budget position, more separating games. The model also has an active-multiplier indicator, cost variance, payoff norm, and payoff correlation. The full model gives AUC 0.835 against 0.795 without the budget-position term, and budget position alone gives 0.742.

D.9.18 Payoff-rescaled cross-K control

After matching the uniform-play gradient variance the peaks remain at 0.95 for $K = 2$ and 0.70 for $K = 32$. Rescaling payoffs by $\sqrt{K/8}$ (uniform-play gradient variance matched to $K = 8$; scale 1 is bitwise the primal-dual learner) leaves the $K = 2$ peak at $\zeta = 0.95$ with height 0.349, against the prediction 0.98 from the geometry of pure strategies and the unscaled measurement 1.0. At $K = 32$ the peak stays at $\zeta = 0.70$ (height 0.851) against the prediction 0.59–0.7. One prediction lies within one grid step of the measured peak and the other contains it.

D.9.19 Dimension profile of the chaotic fraction

Under the growth threshold the chaotic fraction rises smoothly with the number of actions (tangent-map estimate, $G = 500$, $T = 3 \times 10^5$). It is 0.056, 0.128, 0.244 at $K = 3$, 4, 6 at the $K = 8$ peak cell, against 0.32 at $K = 8$. The maxima are 0.120, 0.178, 0.244 when each dimension is instead measured at two cells bracketing its own peak predicted from the geometry of pure strategies. Median exponents of chaotic games are 1–3×10^{-4} per step at every K. A matched classifier audit of the same games (same draws, seeds, and cells) rules out the $K = 2$-style estimator artifact at $K = 3$–6. Among converged games the chaotic fraction is 0.000–0.040 at every small-K cell (fractions at the upper bound 0.00–0.70). Converged-game false-positive rates of 0.5–0.9 had invalidated the $K = 2$ cells with multipliers at the upper bound.

D.9.20 Cross-distribution peak test

The order-statistics reading of the peak of the separating fraction over all games is tested across cost distributions with predictions fixed before the runs. The crossing level is the one calibrated input, taken from the paper: the ζ at which the feasible fraction reaches the 0.882 of the peak cell. Everything else is order statistics of the cost distribution. Our rule predicts the peak at $\zeta \approx 0.90$ for costs uniform on $[0.8, 1.2]$ and ≈ 0.55 for $[0.1, 1.9]$, against 0.749 for the paper's $[0.5, 1.5]$ (measured 0.75). Measurements use fresh 21-cell ζ sweeps with $G = 2{,}000$. At the paper's bounds the game-generation code is bitwise unchanged. The narrow-support peak is at $\zeta = 0.95$ with height 0.728, one grid step from its prediction and a hit against its criterion. The wide-support peak is at $\zeta = 0.45$ with height 0.651, two grid steps from its prediction and a miss against its criterion. A second batch was fully out of sample: its predictions were fixed before the runs, and none of its distributions was used in forming the hypothesis. At $U[0.65, 1.35]$ the prediction is 0.825 and the

Table A20: Exact feasibility split of the aggregate separating fraction along the fine ζ sweep ($\rho = 0$, $\tau = 1$, $T = 3 \times 10^4$, 20,000 games per cell). Feasibility is game-level. The feasible term $= P(\mathrm{f})\,P(\mathrm{dec} \mid \mathrm{f})$ and the infeasible term $= P(\neg\mathrm{f})\,P(\mathrm{dec} \mid \neg\mathrm{f})$ sum to the aggregate (f: feasible, dec: separating). The conv (converged) columns split the converged fraction by feasibility.

ζ	$P(\mathrm{f})$	$P(\mathrm{dec} \mid \mathrm{f})$	$P(\mathrm{dec} \mid \neg\mathrm{f})$	$P(\mathrm{conv} \mid \mathrm{f})$	$P(\mathrm{conv} \mid \neg\mathrm{f})$	feasible term	infeasible term	aggregate
0.5	0.004	0.842	0.066	0.158	0.918	0.003	0.066	0.069
0.55	0.055	0.839	0.220	0.127	0.750	0.046	0.207	0.254
0.6	0.230	0.821	0.385	0.119	0.567	0.189	0.297	0.485
0.65	0.499	0.822	0.495	0.108	0.444	0.410	0.248	0.658
0.7	0.725	0.808	0.557	0.094	0.380	0.586	0.153	0.739
0.75	0.882	0.785	0.599	0.093	0.338	0.692	0.071	0.763
0.8	0.965	0.752	0.563	0.093	0.388	0.725	0.020	0.745
0.85	0.992	0.705	0.556	0.107	0.391	0.700	0.004	0.704
0.9	0.999	0.650	0.750	0.121	0.250	0.650	0.000	0.650
0.95	1.000	0.593	0.000	0.140	n/a	0.593	0.000	0.593
1	1.000	0.536	0.000	0.159	n/a	0.536	0.000	0.536
1.05	1.000	0.460	0.000	0.187	n/a	0.460	0.000	0.460
1.1	1.000	0.398	0.000	0.222	n/a	0.398	0.000	0.398
1.15	1.000	0.331	0.000	0.254	n/a	0.331	0.000	0.331
1.2	1.000	0.268	0.000	0.300	n/a	0.268	0.000	0.268
1.3	1.000	0.155	0.000	0.377	n/a	0.155	0.000	0.155

measurement 0.85 (within one grid step). $U[0.35, 1.65]$ predicted 0.673 and measured 0.65 (within one grid step). A lognormal ($\mu = -0.125$, $\sigma = 0.5$, a non-uniform shape) predicted 0.656 and measured 0.60 (within two grid steps). Across all five held-out distributions the order-statistics rule falls within one grid step three times and within two grid steps twice. Everywhere the peak tracks the feasibility transition in direction and scale. Error grows as the transition broadens (wide supports, heavy shapes).

D.9.21 Feasibility split

Table A20 reports the exact per-cell feasibility split behind the main-text decomposition. At each budget ratio of the fine ζ sweep the aggregate separating fraction equals $P(\mathrm{feas})\,P(\mathrm{dec} \mid \mathrm{feas}) + P(\mathrm{infeas})\,P(\mathrm{dec} \mid \mathrm{infeas})$ by construction, with the feasibility of each game computed exactly from the draws. The feasible-conditioned fraction declines monotonically along the whole axis, up to binomial noise (a $+0.001$ rise at $\zeta = 0.65$ lies inside its 3σ interval). The feasible fraction rises from 0.004 to 1. The aggregate separating fraction peaks at $\zeta = 0.75$, where the feasible and infeasible terms contribute 0.692 and 0.071. That term is zero beyond 0.95. The feasibility split locates the peak. A second-level decomposition ties most of the conditional decline to active multipliers. Splitting feasible games by whether a multiplier is active (final $\max(\lambda_x, \lambda_y) > 0.05$) shows that the fraction with an active multiplier among feasible games falls from 0.96 to 0.21 along the sweep. The separating fraction among feasible games with an active multiplier declines far more gently, from 0.88 at $\zeta = 0.5$ to 0.47 at 1.3. Feasible games without an active multiplier are rarely separating (0.00–0.23). Most of the conditional decline is therefore the decline in the fraction with an active multiplier. The frequency formula and the independent-phase average both require an oscillator with an active multiplier. The converged-fraction columns split feasible from infeasible games. At $\zeta = 0.5$ the converged fraction is 0.918 among infeasible games (trapping at a pure strategy without constraint satisfaction) against 0.158 among feasible ones. Only at looser budgets does feasible convergence grow, from 0.09 to 0.38. Table A21 places the peak of each observable: non-convergence peaks at $\zeta = 0.8$, the separating fraction at 0.75, and the chaotic fraction at 0.75 with 0.8 tied within noise. The feasible term alone peaks at 0.8. In the feasible columns of Table A21 the separating and chaotic fractions fall from $\zeta = 0.6$ on, while the non-converged fraction stays near 0.9 up to $\zeta = 0.8$. Among feasible games at $\zeta = 0.6$–0.9 the most common final configuration is exactly one active multiplier (0.43 to 0.54), and the fraction with both multipliers active falls from 0.53 to 0.23 across these cells.

The geometry organizes the aggregate profile without deciding per-game membership. Static ex ante features (budget positions, cost variances, payoff norms and asymmetry) predict weakly which individual game is chaotic: single-feature AUC is at most 0.55 and joint logistic AUC 0.59 against a base rate 0.34. The same holds for the independent-phase average of the theory. Per game at

Table A21: Peak location by observable along ζ ($\rho = 0$, $\tau = 1$, 20,000 games per cell; chaotic fraction from the long-run ζ sweep, 1,000 games per cell with its own feasible fractions): fractions of all (all) and feasible (f) games, the share whose budget binds ($m < 0$), and the share with one or both multipliers active at the end of the run (threshold 0.05).

ζ	P(f)	non-converged		separating		chaotic		binds | f	one active	both active
		all	f	all	f	all	f			
0.6	0.230	0.536	0.881	0.485	0.821	0.092	0.380	0.988	0.431	0.531
0.7	0.725	0.827	0.906	0.739	0.808	0.282	0.381	0.975	0.487	0.445
0.75	0.882	0.878	0.907	0.763	0.785	0.321	0.362	0.961	0.516	0.387
0.8	0.965	0.896	0.907	0.745	0.752	0.317	0.325	0.939	0.538	0.333
0.9	0.999	0.879	0.879	0.650	0.650	0.307	0.307	0.869	0.544	0.226
1.0	1.000	0.841	0.841	0.536	0.536	0.235	0.235	0.751	0.506	0.150
1.2	1.000	0.700	0.700	0.268	0.268	0.128	0.128	0.394	0.321	0.034

$K = 2$ (the only case where its coefficients are game-level quantities) it distinguishes separating from converged games at AUC 0.55 and correlates with the separation rate at 0.18. The independent-phase average accounts for the aggregate energy growth rather than per-game chaos membership.

Under common random numbers the finite-difference separation rates on the strategy logits have slightly negative medians at every finite budget ratio (-3.8×10^{-6} to -2.0×10^{-9} per step, against $+3.3 \times 10^{-9}$ with no budget). No game is separating.

D.9.22 An independent chaos diagnostic

The Gottwald–Melbourne 0–1 test is run on the same recorded spending series with no reference to any tangent quantity (median K over random frequencies, thresholds fixed before the runs). It agrees with the labels from the tangent-map estimate: median $K_{01} = 0.86$ among chaotic games (criterion > 0.8), 0.01 among cycling games (criterion < 0.2), and 0.01 among the remaining games that are not chaotic (quartiles $[0.00, 0.35]$, consistent with their mostly-regular character). Its stability is checked by random-half resampling of the game population (400 halves). Our K_{01} statistic distinguishes chaotic from other games at AUC 0.883 over 739 usable games, with split-half median 0.883 and 5th–95th percentiles $[0.862, 0.903]$.

D.9.23 Cost-direction-only strategy update

We also intervene on the strategy update. That changes both the regrowth of low-probability actions and the drive of the oscillator. Inside the update each player's logit gradient of the Lagrangian payoff (the payoff field minus the multiplier times the cost vector) is replaced by its Euclidean projection onto the centered and unit-normalized cost vector. The strategy thus moves along the cost direction alone in logit coordinates. The multiplier rule and the spending are computed as before but from the altered strategy trajectory, and their dynamics change with it. The first-order change of the spending $c^\top J_{\mathrm{softmax}}(\ell)\, g$ is not preserved by the projection. The unprojected step is bitwise the main-text step. This collapses the chaotic fraction from 0.31 to 0.034 at the peak cell ($G = 500$, $T = 3 \times 10^5$, tangent-map estimate). The pre-run expectation of a collapse is met. A rerun on the same cell records what else the projection changes, with a matched variant that rescales the projected step so that the first-order change of the spending is preserved exactly. Under the projection the fraction of games whose multiplier is positive over most of the tail falls from 0.75 to 0.47 and the fraction whose spending oscillates in the tail from 0.77 to 0.45. The median tail standard deviation of the spending falls from 0.063 to 0.0001. The spending converges below the budget (median -0.034 against -0.001). The matched variant gives 0.42 with a positive multiplier and 0.08 oscillating and also removes the positive exponents (0.026 against 0.034 under the projection and 0.31 unmodified). Both interventions therefore remove the oscillation itself along with the regrowth of low-probability actions. The projected tangent vectors above and the projected copy below distinguish the regrowth of low-probability actions from the drive.

D.9.24 Cost-direction and orthogonal-complement projections

The subspace-projection measurement propagates sets of tangent vectors along the same base trajectory. Per player and after every tangent step the strategy blocks are projected onto the centered cost direction or its orthogonal complement, with multiplier components kept or zeroed. Identity

Table A22: Projected copy in the long run at the peak cell ($\rho = 0$, $\zeta = 0.75$, $\tau = 1$, $G = 1{,}000$, $T = 3 \times 10^5$, distance 10^{-6} renormalized every 20 steps): chaotic fraction from the finite-difference estimate, median exponent, and share of the chaotic games of the unconfined projected copy. The last two columns repeat the measurement with one tangent vector of the tangent-map estimate projected after every step.

	finite-difference estimate			tangent-map estimate	
deviation confined to	chaotic	median exponent	share	chaotic	share
nothing (unconfined)	0.319	4.3×10^{-6}	1	0.318	1
oscillator directions	0.275	1.8×10^{-6}	0.727	0.277	0.733
other directions, multipliers kept	0.054	1.5×10^{-8}	0.157	0.047	0.142
orthogonal complement (multipliers zeroed)	–	–	–	0.049	0.148

variants reproduce the plain tangent-map estimate and the conditional Lyapunov exponent exactly ($\zeta = 0.75$, $G = 500$, $T = 3 \times 10^5$). Among chaotic games, the orthogonal complement retains a positive conditional exponent in 0.94–0.95 of games at about twice the full exponent (median ratio 2.0, multipliers kept or zeroed alike). The span of the cost direction and the multiplier axes tracks the full exponent (retention 0.91, ratio 1.01) and weakens when its multiplier coupling is set to zero (retention 0.75, ratio 0.83). In this fixed-direction split the growth persists in the orthogonal complement of the constant cost vector, and the span of the cost direction and the multiplier axes sustains positivity mainly through its multiplier coupling. The state-dependent split (the projected tangent vectors above and the projected copy below) shows that growth is generated in the oscillator directions. The expectation stated before the runs, that the cost direction alone would lose positivity, fails: 75% of games retain positivity.

D.9.25 Projected copy

The projected-copy measurement runs two copies of the primal-dual learner on the same game. One copy is never modified, and after every step the projected copy is reset to the unmodified copy plus the projection of their difference onto a subspace. That difference is renormalized to 10^{-6} every 20 steps after $T/2$ and its log growth accumulated ($\zeta = 0.75$, the 1,000 games of the long run at the peak cell, $T = 3 \times 10^5$). This gives a finite-difference estimate of the exponent of the constrained deviation. It labels a game chaotic when the accumulated growth exceeds 3. With no projection the projected copy gives a chaotic fraction of 0.319 against 0.321 for the tangent-map estimate on the same games with per-game label agreement 0.992. With the distance at 10^{-8} the same projected copy gives 0.681, the noise level of the finite-difference estimate (subtraction noise of about 10^{-16} divided by the distance per step). The measurement therefore uses 10^{-6}. Per player the oscillator directions are the state-dependent cost-gradient direction in the logits (components $x_k(c_k - c^\top x)$) plus both multiplier axes. The other directions keep the multiplier deviations and each logit block minus its component along the cost-gradient direction; they are not the orthogonal complement of the oscillator directions. Confined to the oscillator directions the projected copy gives a chaotic fraction of 0.275 (median exponent 1.8×10^{-6}). Confined to the other directions it gives 0.054 (median exponent 1.5×10^{-8}). Among the 319 games that are chaotic under the unconfined projected copy, the projected copy confined to the oscillator directions finds 0.727 chaotic and the one confined to the other directions 0.157 (Table A22). Repeating the measurement with one tangent vector of the tangent-map estimate projected after every step onto the same subspaces gives 0.318, 0.277 (share 0.733), and 0.047 (share 0.142). The strict orthogonal complement of the oscillator directions with the multiplier components zeroed as well gives 0.049, so the multiplier axes add nothing to the growth of the other directions. The two unprojected estimates agree on 0.993 of the labels. The oscillator directions retain the growth while the directions of low-probability actions carry a squared-norm share of 0.87 of the leading Lyapunov vector (Table A6). This agrees with the projected tangent vectors above and with the two interventions on the strategy update. Those interventions remove the oscillation itself.

D.9.26 Mechanism decompositions

At each late-run renormalization a predictor of regrowth records the low-probability action with the largest squared-norm share of the leading Lyapunov vector, its payoff gap (Lagrangian payoff minus the running maximum), and whether its identity changed since the previous renormalization. Across chaotic games the top exponent correlates at 0.47–0.50 with the identity-switching frequency and

Table A23: Chaotic fraction along the long-run ζ sweep, criteria fixed before the runs ($T = 3 \times 10^5$, tangent-map estimate, $G = 1{,}000$ per cell, fresh seeds) with the same-seed classifier separating fraction.

ζ	0.5	0.6	0.7	0.75	0.8	0.9	1.0	1.2
chaotic	0.004	0.09	0.28	0.32	0.32	0.31	0.24	0.13
classifier (same seeds)	0.06	0.47	0.75	0.76	0.75	0.67	0.53	0.27

0.17 with the mean gap depth. The top exponent is thus organized by competition among low-probability actions rather than by how close any one of them is to regrowing. Conditioning on contact with the low-multiplier region sharpens the association. Games that never enter the low-multiplier region post-warm are chaotic at 0.06 and 0.14 (at $\zeta = 0.75$ and 0.8) and games that stay there throughout at 0.05 and 0.08, against 0.30 and 0.34 among games that visit it. Positive exponents concentrate about five-fold among games that visit the low-multiplier region, correlational evidence that complements the interventions. Chaotic games that never visit are few: 5 of 176 and 5 of 202 chaotic games in the two $T = 3 \times 10^4$ cells, about 3%. A same-seed replay shows what they are: their multipliers oscillate around interior levels without ever entering the low-multiplier region or reaching the upper bound (post-warm minimum median 1.2–1.4, mean 4.6–5.7, contact with the upper bound zero, against means of 1.6–2.0 among games that visit). The oscillation driven by the multiplier's integral of past violations is present while the reinjection through the low-multiplier region is not. Visiting the low-multiplier region is the reinjection route for about 97% of chaotic games. The upper bound plays no analogous role. The fraction of games at the upper bound is 0.08 at the peak cell and vanishes above the default upper bound (Table A11). Contact with the upper bound among chaotic games is 0.000–0.005 in the same rerun, and the separating-fraction peak is flat through $\lambda_{\max} = 500$.

The independent-phase average also predicts the ordering across payoff-correlation rows. The competitive coupling statistic of the continuous-time limit, Monte-Carlo averaged over ρ-correlated Gaussian slopes, ranks the measured chaotic fractions across all five ρ rows perfectly (Spearman 1.0 at both $\zeta = 0.75$ and 1.0). It predicts exactly zero at $\rho = +1$, where 10.7% are nonetheless chaotic. The continuous-time limit gives the ordering.

The decomposition of the games that are not chaotic in Appendix D.1 uses the sorted $T = 300{,}000$ Lyapunov spectra of the peak cell from the tangent-map estimate with resolution $\epsilon = 3/150{,}000 = 2 \times 10^{-5}$ per step. The two softmax shift directions contribute one exact zero exponent per player. They are identified in the audits of Appendix D.8 and are not counted toward the multiplicity reported here. Of the 679 games below the growth threshold, every one has both leading exponents inside the resolution interval $|\Lambda| < \epsilon$, so the remainder is neutral at this resolution. Among them 0.91 have Λ_1 positive but below the growth threshold (median 3.7×10^{-6}). The false-positive calibration joins the same Lyapunov spectra to the classifier records. Two of the 99 games the classifier labels converged are chaotic (and 1–9 per cell across the vertex-audit cells). Every one fails argmax stability at its terminal vertex or has a still-moving multiplier under the exact vertex audit. On converged games the classifier and the tangent-map estimate thus agree to within 2%. The chaotic fraction of the long-run ζ sweep has an interior peak at $\zeta = 0.75$ co-located with the classifier separating-fraction peak on the same seeds. Its criteria were fixed before the runs ($T = 300{,}000$, $G = 1{,}000$ per cell, fresh seeds, tangent-map estimate; Table A23). Also requiring the run to be non-converged changes no cell by more than 0.01.

D.9.27 THE (ρ, ζ) CHAOS MAP

The right panel of Figure 2 extends the chaotic fraction to the full (ρ, ζ) grid ($T = 3 \times 10^5$, tangent-map estimate, 400 games per cell, classifier runs paired per game). An interior chaotic-fraction peak exists in every payoff-correlation row except the zero-sum row. In the zero-sum row the chaotic fraction rises monotonically across the scanned range (0, 0.01, 0.66, 0.80, 0.88, 0.96 along ζ) toward the 0.95 of the zero-sum no-budget cell itself. The no-budget column of the panel, measured on fresh draws with the same settings, reads 0.958, 0.225, 0.045, 0.005 and 0 from $\rho = -1$ to $\rho = 1$. That is measured with the same tangent-map estimate and is consistent with the outward-spiral picture of the simultaneous update (Proposition 3). There tighter budgets suppress that pre-existing chaotic fraction and looser ones recover it, with no interior peak. The chaotic-fraction peak moves outward as the payoff correlation decreases: $\zeta = 0.75$ at $\rho \geq 0$ and near 1.0 at -0.5.

D.9.28 FEASIBILITY-CONDITIONED CHAOS MAP

Table A24 re-reads the same grid conditioned on whether each game is feasible (masks recomputed from the deterministic game-generation code at the stored per-cell seeds). The feasible-conditioned fractions sit at or slightly above the overall ones wherever feasibility is partial and decline above the peak in every row, the coarse-grid counterpart of the conditional decline along the fine ζ sweep.

Table A24: Chaotic fraction among feasible games over (ρ, ζ) (tangent-map estimate, $T = 3 \times 10^5$, $G = 400$ per cell; feasible count in parentheses; cells with fewer than 10 feasible games marked n/a).

$\rho \backslash \zeta$	0.25	0.5	0.75	1.0	1.25	1.5
-1.0	n/a	n/a	0.72 (360)	0.80 (400)	0.88 (400)	0.96 (400)
-0.5	n/a	n/a	0.45 (356)	0.47 (400)	0.34 (400)	0.25 (400)
$+0.0$	n/a	n/a	0.28 (348)	0.26 (400)	0.11 (400)	0.06 (400)
$+0.5$	n/a	n/a	0.23 (348)	0.12 (400)	0.03 (400)	0.01 (400)
$+1.0$	n/a	n/a	0.12 (354)	0.07 (400)	0.01 (400)	0.00 (400)

At this per-cell size the 3σ binomial interval is about ± 0.07. The $\rho = 0$ peak cell of the grid gives 0.258, below the $G = 2{,}000$ long-run value 0.338. In that long run 0.889 of the games are feasible, and the chaotic fraction is 0.375 among them and 0.036 among the infeasible ones. A fresh $G = 800$ confirmation cell gives 0.303. The pooled $\rho = 0$ peak-cell estimate is 0.319 over $3{,}200$ games.

Table A25 gives the four-way split on the same paired games. The recall misses (chaotic games labeled cycling) stay between 0 and 0.085 per cell.

At $\rho = 1$, $\zeta = 0.75$ the tangent-map estimate finds 0.107 of games chaotic against 0.000 separating in the common-payoff no-budget cell. The leading Lyapunov vector holds a squared-norm share of 90% on low-probability actions, against 87% at $\rho = 0$. Time-locking to the low-multiplier region is absent there too (growth difference -0.004, correlation -0.02). The chaos at $\rho = 1$ comes from the constraint in games with no payoff conflict, and its leading Lyapunov vector concentrates on the same directions of low-probability actions.

D.9.29 TAIL VIOLATION UNDER DAMPING

The tail-violation audit of damping reruns the $\zeta \in \{0.6, 0.75, 1.2\}$ damping cells with per-game records at $G = 2{,}000$ and matched constants, and reads the tail-violation percentiles (Table A26). The tail violation added by damping is largest in the $\zeta = 0.6$ and 0.75 cells, and only the $\zeta = 1.2$ cell shows a tail improvement over a range of δ.

D.9.30 PROPORTIONAL TERM AGAINST DAMPING

The variant of Section 5.3 adds a proportional term to the multiplier that enters the strategy step, $\lambda_{\text{eff}} = \Pi_{[0,\lambda_{\max}]}(\lambda + k_p\, v)$ with v the current violation. The update of the multiplier itself is unchanged. At a point with zero violation, $v = 0$ and $\lambda_{\text{eff}} = \lambda$, so the term leaves the multiplier used by the strategy unchanged and introduces no feasibility bias. The exploitability of its converged games is 0 from $k_p = 3$ on. Gains were chosen on a held-out cell ($\zeta = 0.7$) and evaluated on the damping-sweep cells ($G = 2{,}000$, $T = 30{,}000$, feasible games; Table A27). The rule takes the largest converged fraction among gains whose tail violation is no larger than that of damping at $\delta = 0.01$ on that cell, and breaks ties by success rate. At $\zeta = 0.75$ the proportional term with $k_p = 30$ converges in 0.756 of games at tail violation 0.0015, while damping reaches that converged fraction only at $\delta = 0.3$ with tail violation 0.173. The tail violation under damping exceeds the undamped 0.032 at every δ, so no damping rate matches the tail violation of the proportional term. The same ordering holds at $\zeta = 0.6$ (0.839 at 0.0018 against 0.772 at 0.251) and at 1.2 (0.526 at 0.0004 against 0.549 at 0.023). At matched converged fraction the tail violation under damping is two orders of magnitude higher. The increase in tail violation that comes with stabilization is thus specific to damping.

The $\zeta = 0.75$ long run at the peak cell confirms the comparison (Table A28). For the Euclidean projection learner, whose row shows a dash, the finite-difference estimate gives a chaotic fraction of 0.006 (Appendix D.3.1). On the games of the long-run ζ sweep at $T = 3 \times 10^5$ the proportional

Table A25: Four-way split at the two central budget ratios of the (ρ, ζ) chaos map ($T = 3 \times 10^4$ classifier paired per game with $T = 3 \times 10^5$ Lyapunov spectra from the tangent-map estimate, 400 games per cell): converged, cycling, separating and chaotic, and separating and not chaotic. Headers abbreviate converged, cycling and separating as conv., cyc. and sep.

	$\zeta = 0.75$				$\zeta = 1.0$			
ρ	conv.	cyc.	sep. and chaotic	sep. and not chaotic	conv.	cyc.	sep. and chaotic	sep. and not chaotic
-1	0.02	0.07	0.65	0.26	0.00	0.14	0.76	0.10
-0.5	0.04	0.10	0.40	0.45	0.06	0.27	0.43	0.23
0	0.12	0.11	0.25	0.51	0.16	0.27	0.25	0.33
0.5	0.20	0.12	0.18	0.49	0.28	0.30	0.11	0.32
1	0.28	0.12	0.10	0.50	0.36	0.21	0.06	0.37

Table A26: Tail-violation audit of damping: tail-violation percentiles (p50/p90/p99) at the undamped baseline, inside the damping range $\delta \in [0.003, 0.03]$ of Table 2, and at full damping ($G = 2{,}000$ per cell).

ζ	$\delta = 0$	$\delta \in [0.003, 0.03]$	$\delta = 1$
0.6	0.040/0.104/0.171	rising monotonically	0.338/0.545/0.704
0.75	0.029/0.066/0.101	p50 dips at $\delta = 0.003$	0.227/0.427/0.582
1.2	n/a/0.038/0.072	improves to n/a/0.018/0.060	0.093/0.198 (p90/p99)

Table A27: Damping against the proportional term on the damping-sweep cells ($\rho = 0$, $\tau = 1$, feasible games, $G = 2{,}000$ per cell): converged fraction / separating fraction / tail violation / 99th-percentile tail violation.

variant	$\zeta = 0.6$	$\zeta = 0.75$	$\zeta = 1.2$
undamped	0.110 / 0.851 / 0.0225 / 0.063	0.085 / 0.788 / 0.0323 / 0.099	0.287 / 0.271 / 0.0107 / 0.072
damping			
$\delta = 0.01$	0.091 / 0.427 / 0.0363 / 0.091	0.123 / 0.427 / 0.0328 / 0.103	0.324 / 0.176 / 0.0062 / 0.061
$\delta = 0.03$	0.611 / 0.273 / 0.0704 / 0.162	0.547 / 0.291 / 0.0526 / 0.155	0.492 / 0.137 / 0.0085 / 0.059
$\delta = 0.3$	0.772 / 0.056 / 0.2512 / 0.586	0.755 / 0.050 / 0.1728 / 0.468	0.549 / 0.056 / 0.0233 / 0.167
proportional term			
$k_p = 3$	0.513 / 0.364 / 0.0062 / 0.046	0.505 / 0.355 / 0.0073 / 0.057	0.508 / 0.110 / 0.0014 / 0.027
$k_p = 10$	0.718 / 0.184 / 0.0034 / 0.039	0.663 / 0.201 / 0.0037 / 0.040	0.533 / 0.073 / 0.0007 / 0.013
$k_p = 30$	0.839 / 0.077 / 0.0018 / 0.023	0.756 / 0.104 / 0.0015 / 0.024	0.526 / 0.066 / 0.0004 / 0.007

term with $k_p = 30$ is chaotic on 0.012 of games against 0.321 undamped. Of feasible games 0.80 end converged within 0.005 of the budget and within 0.02 of a constrained equilibrium. Damping at $\delta = 0.3$ reaches a converged fraction of 0.81 only with tail violation 0.18. The success rate of damping is then 0.03.

The optimistic and extragradient updates are tuned by the same rule on the same held-out cell over steps from 0.01 to 4. Their one free parameter is the step, since the extrapolation coefficient is one and the half step uses the same step. The rule selects $\eta = 1$ for optimism and $\eta = 2$ for extragradient. At those steps on the same games optimism reaches success rate 0.698 with chaotic fraction 0.115 and extragradient 0.656 with 0.168. The proportional term has the higher success rate at every step of the sweep. The Euclidean and KL projection learners meet the budget and reach success rates 0.734 and 0.741. Their other feasible games cycle with a negative exponent. Among the tested remedies the proportional term therefore has the highest success rate and the projection learners follow when the feasible set is known. The tuned optimism and extragradient updates and damping trail behind. The criteria were stated before the runs, and the projection learners run at the paper's strategy step. The proportional term leads the projections by 0.061 ± 0.027 (KL) and 0.067 ± 0.027 (Euclidean), paired 95% intervals on the same 875 feasible games.

D.9.31 VERTEX AUDIT AND HORIZON STABILITY

The vertex audit reruns the three cells $\zeta = 0.5$, 0.6, 0.8 on the seeds of the long-run ζ sweep and applies an exact stability test to every game the classifier labels converged (all of which end at vertices). The occupied action pair is checked against the argmax of the exact Lagrangian payoff at the terminal state. This decision does not depend on floating-point precision: the logit of a low-probability action moves at rate ηq_k whatever its probability. The audited property is exact: a

Table A28: Remedies in the long run at the peak cell ($\rho = 0$, $\zeta = 0.75$, $\tau = 1$, $G = 1{,}000$, $T = 3 \times 10^5$; tangent-map estimate; feasible games): chaotic fraction (all / feasible), converged fraction, success rate (converged, tail violation at most 0.005, last-iterate exploitability at most 0.02), tail violation and 99th-percentile tail violation, median last-iterate exploitability, tail multiplier. The two projection learners have no multiplier.

variant	chaotic	converged	success rate	tail violation	p99 tail violation	exploitability (last)	tail λ
undamped	0.321 / 0.362	0.038	0.037	0.0335	0.096	0.199	3.02
damping							
$\delta = 0.01$	0.165 / 0.175	0.558	0.104	0.0324	0.096	0.000	2.01
$\delta = 0.3$	0.019 / 0.021	0.809	0.031	0.181	0.475	0.000	0.60
proportional term							
$k_p = 10$	0.020 / 0.023	0.763	0.741	0.0028	0.037	0.000	2.97
$k_p = 30$	0.012 / 0.014	0.825	0.801	0.0009	0.017	0.000	2.94
optimism, tuned $\eta = 1$	0.115 / 0.128	0.707	0.698	0.0071	0.068	0.000	2.99
extragradient, tuned $\eta = 2$	0.168 / 0.189	0.669	0.656	0.0086	0.070	0.000	2.98
Euclidean projection	–	0.734	0.734	0	0	0.000	–
KL projection	0.012 / –	0.750	0.741	0	0	0.000	–

Table A29: Constraint-count controls at $\zeta = 0.7$ / 0.75 / 0.8, same drawn games: separating fractions, multipliers at the upper bound, and the feasibility-conditioned gap.

Measurement	Values
Separating at $\zeta = 0.7/0.75/0.8$, for single; duplicated; independent	0.74/0.77/0.74; 0.73/0.73/0.71; 0.89/0.93/0.93
At least one multiplier at the upper bound	0.31–0.83 (independent) vs 0.02–0.20 (others)
Independent minus duplicated, feasibility-conditioned	0.11, 0.15, 0.18
Independent-variant game feasibility	0.07, 0.21, 0.48 (vs 0.72–0.96 in the other variants)

vertex pair is a fixed point of the argmax dynamics on the Lagrangian payoff if and only if $\lambda_{\max} \cdot (c_k - c_{\min}) \geq u_k - u_{\min}$ for each player and every action k against the opponent's vertex. With unbounded Gaussian payoffs the condition fails with positive probability (the non-trapped remainder below). The argmax-stable fraction is 0.95, 0.79, 0.66 at $\zeta = 0.5$, 0.6, 0.8. The games that are not argmax-stable are slow transients whose escape-time estimates (logit gap over η times payoff gap) have medians 16,000–24,000 steps, of order the horizon. These transients match the slow erosion of trapping under longer runs (converged fraction $0.92 \to 0.90$ at $\zeta = 0.5$ and $0.50 \to 0.43$ at 0.6 from $T = 30{,}000$ to $120{,}000$). Under the 300,000-step Lyapunov spectra from the tangent-map estimate the few chaotic converged games (1, 6, 9 at $\zeta = 0.5$, 0.6, 0.8) all fail argmax stability at their terminal vertex or have still-moving multipliers. That is the slow-transient subpopulation the audit identifies. No game converged at an argmax-stable vertex is chaotic. The trapped fraction stabilizes: 0.92, 0.90, 0.90 at $\zeta = 0.5$ and 0.50, 0.43, 0.40 at 0.6 for $T = 3 \times 10^4$, 1.2×10^5, 3×10^5, with the erosion decelerating.

D.9.32 Constraint-count controls

The duplicated-budget control (Table A29; construction in Appendix D.7) keeps the two multipliers identical. The duplicated-budget control stays at the single-budget level while the variant with two independent budgets rises. Among feasible games an independent second constraint raises the separating fraction by 0.11–0.18 over the duplicated budget, so matched gain alone does not produce the amplification.

The joint-action-cost profile (Table A30) passes its test for an interior peak, set before the runs, in both the separating and non-converged forms. Vertex trapping is weaker than under own-strategy costs, because a feasible vertex for one player now depends on the opponent's strategy.

D.9.33 Comparison of the two updates

Table A31 is the comparison of the two updates behind Section 4.1: five cells along the budget ratio rerun under the simultaneous update on the same drawn games and seeds as their multiplier-first counterparts ($G = 2{,}000$ per cell). The no-budget cell is bit-identical between the two updates, as

Table A30: Joint-action-cost profile ($G = 2{,}000$ per cell): separating and converged fractions along the budget ratio and in the no-budget cell.

ζ	0.5	0.6	0.7	0.75	0.8	0.9	1.0	1.2	∞
separating	0.06	0.44	0.69	0.69	0.71	0.71	0.60	0.35	0.03
converged	0.54	0.35	0.23	0.23	0.19	0.17	0.18	0.26	0.51

they differ only through the multiplier. The expectation stated before the runs was that the simultaneous update would show a larger separating fraction, because the linearized simultaneous update is uniformly weakly expanding. It is not borne out. The effect of switching updates is negligible to slightly stabilizing.

Table A31: Comparison of the two updates at $\rho = 0$, $\tau = 1$, 2,000 games per cell, matched draws and seeds: fractions of classifier labels (converged/cycling/separating) under the multiplier-first and simultaneous (theorem) updates.

ζ	multiplier-first	simultaneous
0.6	.476/.047/.478	.477/.048/.475
0.75	.122/.113/.766	.126/.117/.758
0.9	.122/.228/.651	.123/.240/.638
1.2	.278/.445/.278	.277/.480/.244
∞	.540/.442/.019	.540/.442/.019

D.9.34 Markov-game record

Table A32 is the per-cell record of the constrained Markov game variant. Its tight-budget cells are re-measured at ten times the default upper bound per the standing protocol: convergence 0.93, 0.85, 0.34 at $\zeta = 0.5, 0.6, 0.7$ against 0.94, 0.87, 0.35 at the default upper bound, all within sampling intervals. The cycling-separating split shifts exactly as in the one-shot tight-budget cells. Its no-budget cell is itself 0.19 separating, and the excess at the peak there is quoted as the absolute rise over the no-budget value.

Table A32: Constrained Markov game variant (S=4, K=4, $\rho = 0$, $\tau = 1$, discount 0.9, 2,000 games per cell): converged, cycling and separating fractions along the budget ratio.

ζ	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.5	∞
converged	0.94	0.87	0.35	0.06	0.03	0.07	0.14	0.18	0.21	0.21
cycling	0.06	0.07	0.07	0.07	0.12	0.27	0.43	0.56	0.61	0.60
separating	0.00	0.07	0.59	0.87	0.84	0.66	0.43	0.26	0.19	0.19

D.10 Per-cell records

Table A34 is the coarse grid behind the 43-of-45 count of Section 5.1, distinct from the 20,000-game cells behind Figure 2. The fine ζ sweep behind Figure 3 and the damping sweep behind Figure A5 are per-cell records in the supplementary material. Table A36 consolidates the robustness-set verdicts with the generic robustness variants and the protocol audits. Figure A4 shows one representative trajectory for each of three behaviors: absorption at a feasible pure strategy, a separating run near the peak, and a limit cycle at a large step-size ratio. Figure 1 shows a two-action game ($\rho = 0$, $\zeta = 1$, $\tau = 1$) with competitive coupling and both multipliers positive for part of the tail. Its run approaches a periodic oscillation. All cells: $\eta = 0.05$, $T = 30{,}000$, $K = 8$, $\lambda_{\max} = 50$ unless stated. Every $\rho \neq 0$ ensemble uses the payoff-correlation construction of Appendix B.

D.10.1 Contact with zero and the upper bound

Table A35 records the fractions behind the scope statement of Section 4.2. In the peak cell 79% of the games that the classifier labels as separating touch zero within the classifier run. Of all games 98% touch zero or the upper bound by 3×10^5 steps (zero alone: 93%). Contact is recorded after the 24,000-step warm-up with the same draws, seeds, and program as the classifier cell.

Table A33: Classifier threshold audit at $\rho = 0$, $\tau = 1$, 20,000 games per cell: each entry is the separating-fraction peak (separating fraction at its interior peak ζ) over the no-budget separating fraction under the same thresholds.

conv. thr. \ rate thr.	10^{-2}	10^{-3}	10^{-4}
10^{-5}	0.480 @ 0.70 / 0.000	0.762 @ 0.75 / 0.027	0.881 @ 0.80 / 0.468
10^{-6}	0.481 @ 0.70 / 0.000	0.763 @ 0.75 / 0.027	0.882 @ 0.80 / 0.472
10^{-7}	0.482 @ 0.70 / 0.000	0.764 @ 0.75 / 0.027	0.883 @ 0.80 / 0.476

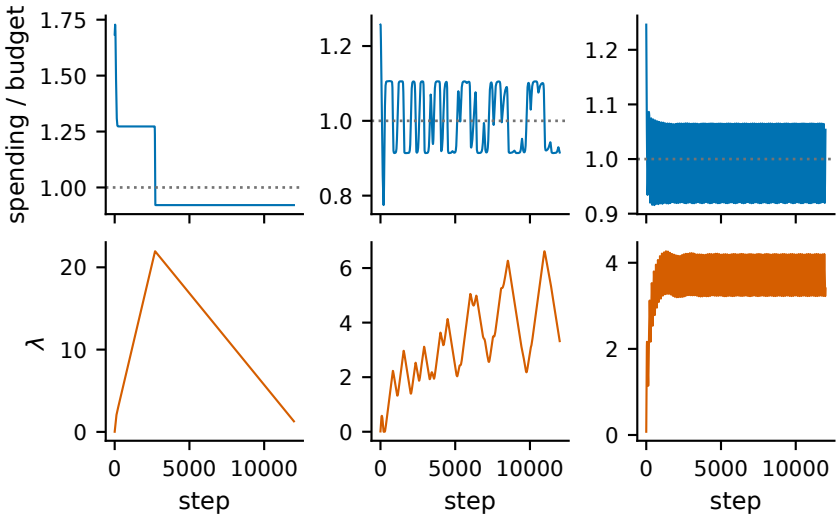


Figure A4: Spending relative to budget (top) and multiplier (bottom) for representative games at $\rho = 0$. Left to right: absorption at a feasible pure strategy ($\zeta = 0.6$), separating run near the peak ($\zeta = 0.8$), limit cycle at a large step-size ratio ($\zeta = 0.8$, $\tau = 10$).

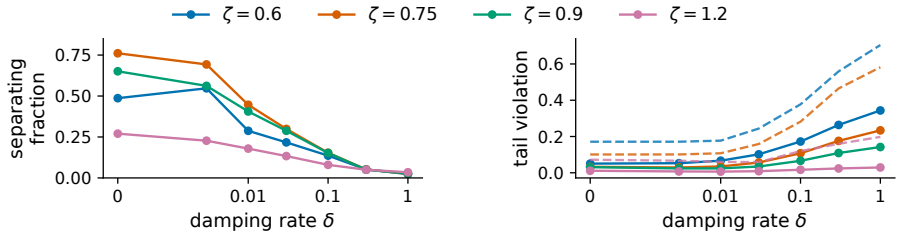


Figure A5: Beyond $\delta = 0.03$ damping lowers the separating fraction at every budget ratio and raises tail violation. The smallest rates can reverse this: at $\zeta = 0.6$, $\delta = 0.003$ raises the separating fraction from 0.49 to 0.55 and lowers tail violation in most cells. Left: separating fraction, δ axis linear near zero, logarithmic beyond. Right: tail violation, mean (solid) and 99th percentile (dashed).

D.10.2 Welfare of chaotic games under damping

A same-seed pair at the peak cell ($G = 500$, $T = 3 \times 10^5$; damped dynamics and tangent-map estimates on identical draws) conditions the equilibrium-quality records on whether each game is chaotic. Damping at $\delta = 0.01$ lowers the chaotic fraction from 0.318 to 0.138 and raises the converged fraction from 0.054 to 0.624. The remaining chaotic games are no closer to equilibrium: median last-iterate exploitability 0.475 against 0.445 undamped, with tail violation 0.048 against 0.034. Tail-averaged exploitability improves (0.027 to 0.005). That remainder stays far from equilibrium.

D.10.3 Multiplier initialization

Initializing both multipliers at the median active multiplier $\lambda_0 = 4.3$ instead of zero leaves the peak-cell chaotic fraction unchanged within noise (0.318 vs 0.321, $G = 500$ vs the $G = 1{,}000$ horizon-sequence value, fraction-level comparison since the draw streams differ).

Table A34: Coarse grid: outcome fractions (converged/cycling/separating) over (ρ, ζ, τ), 2,000 games per cell. The ∞ rows are τ-independent because the multipliers never activate, and they span the three τ columns.

ρ	ζ	τ=0.1	τ=1	τ=10
−1.0	0.5	0.816/0.004/0.180	0.887/0.029/0.084	0.876/0.051/0.073
−1.0	0.8	0.006/0.017/0.977	0.004/0.059/0.936	0.006/0.315/0.678
−1.0	1.2	0.001/0.060/0.939	0.001/0.152/0.847	0.002/0.301/0.697
−1.0	∞		0.002/0.076/0.922	
−0.5	0.5	0.867/0.003/0.131	0.907/0.018/0.074	0.887/0.045/0.068
−0.5	0.8	0.113/0.067/0.821	0.037/0.119/0.844	0.022/0.335/0.642
−0.5	1.2	0.127/0.409/0.464	0.116/0.416/0.468	0.104/0.516/0.380
−0.5	∞		0.223/0.588/0.189	
+0.0	0.5	0.895/0.003/0.102	0.902/0.015/0.082	0.897/0.037/0.065
+0.0	0.8	0.255/0.100/0.645	0.107/0.155/0.738	0.083/0.410/0.506
+0.0	1.2	0.335/0.395/0.271	0.285/0.435/0.279	0.289/0.517/0.194
+0.0	∞		0.502/0.466/0.032	
+0.5	0.5	0.910/0.003/0.087	0.926/0.017/0.058	0.903/0.033/0.064
+0.5	0.8	0.393/0.099/0.509	0.197/0.148/0.655	0.155/0.420/0.425
+0.5	1.2	0.535/0.260/0.205	0.480/0.322/0.198	0.454/0.419/0.128
+0.5	∞		0.799/0.200/0.001	
+1.0	0.5	0.920/0.006/0.073	0.921/0.010/0.069	0.893/0.033/0.074
+1.0	0.8	0.509/0.098/0.393	0.276/0.165/0.559	0.265/0.375/0.360
+1.0	1.2	0.632/0.208/0.160	0.566/0.222/0.212	0.565/0.296/0.139
+1.0	∞		0.995/0.005/0.000	

Table A35: Contact with zero and the upper bound in the peak cell ($K = 8$, $\rho = 0$, $\zeta = 0.75$, $\tau = 1$, $G = 2{,}000$, $\eta = 0.05$, $\lambda_{\max} = 50$): fraction of games whose multiplier touches the bound $\lambda = 0$, or that bound or the upper bound, by each horizon. Separating = the 1,506 games that the classifier labels as separating in its standard 30,000-step run.

	zero		zero or upper bound	
Horizon	all games	separating games	all games	separating games
3×10^4	0.813	0.788	0.862	0.828
10^5	0.908	0.902	0.962	0.952
3×10^5	0.927	0.926	0.981	0.977

D.10.4 Lyapunov spectra with more players

Tangent-map estimates at the N-variant peak cells (Table A18) confirm the passes above with Lyapunov exponents. The many-player chaotic fractions peak in the interior and exceed the two-player peak value, while the no-budget values stay at the two-player no-budget level. The excess is thus constraint-made rather than generic many-player chaos. Markov-variant tangent-map estimates follow the same pattern: 0.005/0.605/0.33/0.29 at $\zeta = 0.5/0.8/1.2/\infty$ (S=4, K=4, $G = 200$) with an interior peak of the chaotic fraction at its classifier peak and the no-budget value well below it.

Table A36: Robustness set and protocol audits: verdicts against criteria set before the runs. Peak = interior peak beyond 3σ above the values at ζ=0.5, 1.2, and ∞ (chaotic fractions first, classifier readings in parentheses); tightest cell = converged fraction at the tightest budget ratio, invariant to the upper bound. Mechanism controls: Table 2; variant constructions: Appendix D.7.

What changed	Peak height and verdict	Tightest cell
actions K=2 / K=32	non-converged-fraction form at K=2 (peak 0.39; quasi-periodic) / (classifier peak 0.84 at ζ=0.6–0.7): kept	0.999 / 0.797 at upper bound ×10
payoffs t_3/uniform/sparse/rank-2	peak 0.28/0.29/0.19/0.25 vs own no-budget value $\leq$ 0.05: kept (classifier peaks 0.75 at ζ=0.75 / 0.76 at 0.7 / 0.60 at 0.75 / 0.72 at 0.75)	0.907 / 0.915 / 0.934 / 0.906 at upper bound ×10
projected-gradient update	classifier peak in both test forms, cycling-shifted: kept	0.888 at upper bound ×10
optimistic / extragradient (η swept)	peak 0.07–0.24 vs flanks 0.01–0.09: kept, thins at large steps (classifier peaks 0.69 at ζ=0.7 / 0.74 at 0.75)	0.912 / 0.914 at upper bound ×10
step size fourfold (η=0.02 / 0.1), $\eta^2 T$ fixed	within ±0.05 of the cell mean, peak fixed: collapses onto $\eta^2 T$ (classifier peaks 0.71 at ζ=0.75 / 0.77 at 0.7–0.8)	0.916 / 0.902 at upper bound ×10
two budgets (M=2), independent / duplicated	independent 0.45/0.63 vs one budget 0.32/0.31; duplicated 0.34/0.28 at the one-budget level: amplified by a second independent multiplier rather than by gain (classifier peak 0.93)	reformulated: convergence to a feasible pure strategy
joint-action costs $C(i, j)$	0.28 vs flanks 0.09/0.15: kept (classifier peak 0.71 at ζ=0.8–0.9, both forms)	weakened: feasible pure strategy depends on opponent
$N = 3$ (K=8) / 5 (K=4) players	0.64 / 0.65 vs no-budget values 0.04 / 0.06: kept (classifier peaks 0.89 at ζ=0.8 / 0.91 at 1.0)	0.877 / 0.897 at upper bound ×10
Markov game (S=4, K=4)	0.61 vs no-budget value 0.29: kept (classifier peak 0.87 at ζ=0.8; no-budget value 0.19)	0.933 at upper bound ×10
payoff-correlation rows (ζ=0.75, feasible games)	0.72 / 0.45 / 0.28 / 0.23 / 0.12 at $\rho = -1$ to $+1$ vs looser ζ=1.5: 0.96 / 0.25 / 0.06 / 0.01 / 0.00: kept except the zero-sum row (Table A24)	not applicable
simultaneous update	(maximum classifier deviation 0.035): immaterial	not applicable
thresholds (nine) / upper bound 12.5–500	peak position 0.70–0.80 / (classifier 0.69–0.77): unchanged	not applicable
initializations, burn-in, estimator	0.317–0.322 / 0.343 → 0.292 / labels agree 0.97–0.99: burn-in only	not applicable