CONSTANT INDIVIDUAL REGRET IN GENERAL GAMES

Anonymous authors
Paper under double-blind review

ABSTRACT

Uncoupled no-regret dynamics provide a decentralized route to equilibrium. We study horizon-independent individual regret in finite N-player normal-form games under full-information feedback. We show that standard Optimistic Hedge has worst-case regret at least exponential in either the number of players N or the largest action-set size $m_{\max}$. This motivates *ECHO-OFTRL*: optimistic follow-the-regularized-leader (OFTRL) equipped with an EMA cascade for high-order optimism (ECHO), where EMA denotes exponential moving average. The algorithm is deterministic and fully uncoupled. Simultaneously, for every horizon $T \geq 1$, each player in self-play incurs regret upper bounded by $O(N \log m_{\max})$. With a regret-dependent learning rate, it also guarantees regret bounded by the self-play bound plus $O(\sqrt{T \log m_{\max}})$ in the adversarial setting. More generally, we relate the regret bound to the order of the prediction-error filter's zero at one.

1 INTRODUCTION

Uncoupled no-regret dynamics model decentralized learning: each player updates from its own full payoff vector without knowing the other players' objectives (Hart & Mas-Colell, 2000; Greenwald & Jafari, 2003). If every player has small external regret, empirical play approaches the set of coarse correlated equilibria (CCE) (Moulin & Vial, 1978; Cesa-Bianchi & Lugosi, 2006). Regret-minimization methods also underlie several landmark systems for solving large imperfect-information games (Bowling et al., 2015; Moravčík et al., 2017; Brown & Sandholm, 2018; 2019).

Traditional online-learning analysis treats the payoff sequence as an arbitrary, possibly adversarial, exogenous process. Against such a sequence, $\sqrt{T}$ regret can be easily shown to be unavoidable (Hazan, 2016). In self-play, however, the payoff sequences are coupled through a fixed game and the learners' updates. While nonstationary, the smooth update of strategies makes the dynamic landscape faced by each player in the game slightly predictable. Optimistic methods seek to exploit this predictability to accelerate learning and convergence to equilibrium. Optimism has also been studied more broadly as a route to acceleration in optimization and game solving (Wang & Abernethy, 2018; Mertikopoulos et al., 2019; Farina et al., 2019; Piliouras et al., 2022a). Fast rates under robustness and nonparametric structure have also been studied (Foster et al., 2016; Daskalakis & Golowich, 2022).

The quest for accelerated learning dynamics has driven a long progression of faster rates. Predictable-sequence methods brought optimism to online learning (Chiang et al., 2012; Rakhlin & Sridharan, 2013). The RVU framework (Syrgkanis et al., 2015) then gave $T^{1/4}$ individual regret and, more crucially, exposed a key mechanism of analysis that has since been a key mainstay. Unfortunately, that analysis alone is capable of guaranteeing constant *sum* of regrets, but not constant individual regret, as negative regrets can hide another player's large positive regret.

Later, Chen & Peng (2020) leveraged the RVU bound, but was able to break through the $T^{1/4}$ barrier in the special case of two-player games. Breaking past the polynomial dependence in T, the high-order, discrete-time frequency analysis of Optimistic Hedge of Daskalakis et al. (2021) reduced individual regret to $O(\log^4 T)$. This approach was subsequently extended to certain structured polyhedral games (Farina et al., 2022b). Lifted OFTRL then obtained $O(\log T)$ regret over general convex games by turning ordinary regret into a nonnegative lifted regret (Farina et al., 2022a). Dynamic learning-rate control matched this logarithmic horizon dependence while improving the action dependence to polylogarithmic (Soleymani et al., 2025b). This mechanism proved extremely general and robust across different choices of the convex action set and regularizer (Soleymani et al.,

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.

Table 1: Progress toward horizon-independent individual regret in self-play via uncoupled learning dynamics.

Method	Individual regret	Main technique(s)
OFTRL/OOMD (Syrgkanis et al., 2015)	$O(\sqrt{N}\log m_{\max} T^{1/4})$	One-step optimism and RVU bounds
Optimistic Hedge (Chen & Peng, 2020)	$O\big(\log^{5/6} m_{\max} T^{1/6}\big)$ (Two-players only)	Coupled strategy/payoff path-length bounds
Optimistic Hedge (Daskalakis et al., 2021)	$O\big(N \log m_{\max} \log^4 T\big)$	Analysis leverages high-order discrete differences
LRL-OFTRL (Farina et al., 2022a)	$O(N m_{\max} \log T)$	Lifting, nonnegative RVU bounds, and log regularization
DLRC-OMWU / Cautious Optimism (Soleymani et al., 2025b;a)	$O\big(N \log^2 m_{\max} \log T\big)$	Dynamic pacing, intrinsic Lipschitzness
ECHO-OFTRL (This paper)	$O\big(N \log m_{\max}\big)$	Transfer-function-designed EMA residual and nonnegative RVU bound

2025a). Related work also obtains logarithmic swap regret in multiplayer games (Anagnostides et al., 2022). Finally, Clairvoyant MWU (Piliouras et al., 2022b) obtained a constant regret bound from a fixed-point update that evaluates payoff vectors at the strategy profile being computed. This guarantee is not directly comparable to those above: the method requires unrolling joint fixed-point computations, and its regret bound applies only to a sparse subsequence rather than the full history of play.

We summarize the comparable results in Table 1 and a detailed discussion of concurrent work is postponed to Section A. In Table 1, N denotes the number of players and $m_{\max}$ denotes the largest number of actions of any player. All bounds are specialized to utilities in $[0, 1]$; the row for OFTRL and optimistic online mirror descent (OOMD) uses entropy regularization. This history points to a natural question:

> *Can uncoupled regularized learning keep every player's positive regret uniformly bounded in time?*

We show that it can for full-information self-play in finite normal-form games, resolving the horizon-dependence question in this setting.

1.1 Contributions

We first show that the horizon-independent regret bound for Optimistic Hedge established by Ha (2026a) must depend exponentially on either the number of players N or the largest action-set size $m_{\max}$ in the worst case.

Theorem 1.1 (Lower bound)**.** *Consider finite normal-form games with utilities in* $[0, 1]$ *and full-information feedback. For Optimistic Hedge initialized at the uniform strategies, the worst-case individual regret is at least* $2^{\Omega(N)}$ *in an* N*-player game with eight actions per player, and at least* $2^{\Omega(m)}$ *in a six-player game with* m *actions per player.*

These lower bounds show that exponential dependence on either N or $m_{\max}$ is unavoidable even when the other parameter is held fixed. The formal statements are given in Theorems 3.1 and 3.2. This does not contradict the $O(N \log m_{\max} \log^4 T)$ upper bound of Daskalakis et al. (2021), since that bound still grows unboundedly with T.

We then introduce ECHO-OFTRL, which improves the dependence on the number of players to linear and on the number of actions to logarithmic.

Theorem 1.2 (ECHO-OFTRL, informal)**.** *For every* N*-player finite normal-form game with utilities in* $[0, 1]$ *and full-information feedback, ECHO-OFTRL is deterministic and uncoupled, has a closed-*

form update, and uses a common constant learning rate depending only on N. When all players use it, every player $i \in [N]$ satisfies

$$\mathrm{Reg}_i(T) \le O\left(N \log m_{\max}\right)$$

simultaneously for every horizon $T \ge 1$.

The formal statement appears in Theorem 4.2. In Section 5, we extend the analysis to general prediction-error filters. Roughly, using a kth-order predictor with a suitable learning rate gives an individual regret bound with horizon dependence $T^{1/(2k+2)}$ for $0 \le k < N$, while sufficiently high-order prediction gives a horizon-independent bound at a constant learning rate; the precise conditions are stated in Theorem 5.2. Figure 1 illustrates the horizon dependence. This recovers the $\sqrt{T}$ dependence of FTRL without optimism ($k = 0$) (Hazan, 2016), the $T^{1/4}$ dependence of the RVU bound with one-step optimism ($k = 1$) (Syrgkanis et al., 2015), and includes the horizon-independent bound of ECHO-OFTRL.

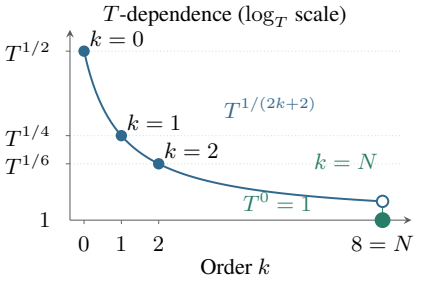

Figure 1: Horizon dependence in Theorem 5.2 for $N = 8$.

2 Preliminaries

For every positive integer n, let $[n] \coloneqq \{1, \dots, n\}$. For a vector $\boldsymbol{z} \in \mathbb{R}^n$ and $a \in [n]$, write $z(a)$ for its ath coordinate. For $1 \le p < \infty$, define $\|\boldsymbol{z}\|_p \coloneqq \left(\sum_{a=1}^n |z(a)|^p\right)^{1/p}$, and define $\|\boldsymbol{z}\|_\infty \coloneqq \max_{a \in [n]} |z(a)|$. For a scalar $r \in \mathbb{R}$, define $[r]_+ \coloneqq \max\{r, 0\}$. The $(n-1)$-dimensional probability simplex is $\Delta^n \coloneqq \{\boldsymbol{x} \in [0,1]^n : \sum_{a=1}^n x(a) = 1\}$.

We use $\langle \cdot, \cdot \rangle$ for the standard inner product and $\mathbf{1}$ for the all-ones vector of the appropriate dimension. For a statement P, set $\mathbb{1}_{\{P\}} = 1$ if P holds and $\mathbb{1}_{\{P\}} = 0$ otherwise. For a tuple $\boldsymbol{x} \coloneqq (\boldsymbol{x}_i)_i$, write $\boldsymbol{x}_{-i} \coloneqq (\boldsymbol{x}_j)_{j \ne i}$. Superscripts in parentheses index time, so $\boldsymbol{z}^{(t)}$ denotes the value of a sequence $\boldsymbol{z}$ at timestep t. For a differentiable map $\Phi : E \to F$ between finite-dimensional Euclidean spaces, $\boldsymbol{x} \in E$, and $\boldsymbol{\xi} \in E$, $\boldsymbol{J}_\Phi(\boldsymbol{x}) : E \to F$ denotes its Jacobian at $\boldsymbol{x}$, and $\boldsymbol{J}_\Phi(\boldsymbol{x})[\boldsymbol{\xi}] \coloneqq \frac{\mathrm{d}}{\mathrm{d}r} \Phi(\boldsymbol{x} + r\boldsymbol{\xi})\big|_{r=0}$ is its directional derivative along $\boldsymbol{\xi}$. If $E = \prod_{j=1}^k E_j$, then, for $j \in [k]$, write $\partial_j \Phi$ for the derivative with respect to the jth block. For scalar-valued f, ∇f denotes its gradient, so $\boldsymbol{J}_f(\boldsymbol{x})[\boldsymbol{\xi}] = \langle \nabla f(\boldsymbol{x}), \boldsymbol{\xi} \rangle$, and $\nabla^2 f$ denotes its Hessian.

2.1 Finite normal-form games and regret

There are $N \ge 2$ players. Player i has action set $[m_i]$, where $m_i \ge 2$, and chooses a mixed strategy $\boldsymbol{x}_i \in \Delta^{m_i}$. Define the largest action-set size by

$$m_{\max} \coloneqq \max_{i \in [N]} m_i.$$

Player i's utility function $\mathcal{U}_i \colon \prod_{j=1}^N \Delta^{m_j} \to [0,1]$ is the multilinear extension of a pure-profile payoff function taking values in $[0,1]$. Define the expected payoff vector by

$$\mathcal{U}_i(\boldsymbol{x}_{-i}) \coloneqq \nabla_{\boldsymbol{x}_i} \mathcal{U}_i(\boldsymbol{x}) \in [0,1]^{m_i}.$$

It is multiaffine in the opponents' strategies and does not depend on $\boldsymbol{x}_i$.

For a horizon $T \ge 1$, at every timestep $t \in [T]$, the players choose $\boldsymbol{x}^{(t)} \in \prod_{j=1}^N \Delta^{m_j}$ simultaneously and player i observes the full vector $\boldsymbol{u}_i^{(t)} \coloneqq \mathcal{U}_i(\boldsymbol{x}_{-i}^{(t)})$. Its external regret is

$$\mathrm{Reg}_i(T) \coloneqq \max_{\widehat{\boldsymbol{x}}_i \in \Delta^{m_i}} \sum_{t=1}^T \left[\mathcal{U}_i(\widehat{\boldsymbol{x}}_i, \boldsymbol{x}_{-i}^{(t)}) - \mathcal{U}_i(\boldsymbol{x}^{(t)})\right] = \max_{\widehat{\boldsymbol{x}}_i \in \Delta^{m_i}} \sum_{t=1}^T \left\langle \boldsymbol{u}_i^{(t)}, \widehat{\boldsymbol{x}}_i - \boldsymbol{x}_i^{(t)} \right\rangle.$$

The equality follows from multilinearity in player i's strategy.

3 OPTIMISTIC HEDGE

Optimistic Hedge uses the preceding payoff vector to predict the next one. We quantify the worst-case dependence of its regret on the number of players and actions.

Fix a learning rate $\eta > 0$. For every player $i \in [N]$, Optimistic Hedge starts from $x_i^{(1)}(a_i) = 1/m_i$ for every $a_i \in [m_i]$ and $\boldsymbol{u}_i^{(0)} = \mathbf{0}$. At every timestep $t \geq 1$, its next strategy satisfies the following update for every $a_i \in [m_i]$:

$$x_i^{(t+1)}(a_i) = \frac{x_i^{(t)}(a_i) \exp\left(\eta\left(2u_i^{(t)}(a_i) - u_i^{(t-1)}(a_i)\right)\right)}{\sum_{b_i \in [m_i]} x_i^{(t)}(b_i) \exp\left(\eta\left(2u_i^{(t)}(b_i) - u_i^{(t-1)}(b_i)\right)\right)}. \qquad \text{(Optimistic Hedge)}$$

By Ha (2026a, Theorem 4.1), a sufficiently small common constant learning rate gives constant individual regret, with an implicit dependence on the game size. The next two results show that this dependence must be at least exponential in either the number of players or the number of actions, even when the other parameter is fixed.

3.1 DEPENDENCE ON THE NUMBER OF PLAYERS

Write $\mathrm{Reg}_i^G(T)$ for player i's external regret in game G under (Optimistic Hedge). In the following bounds, all players use a common constant learning rate $\eta > 0$, and the supremum over T ranges over positive integer horizons.

Theorem 3.1 (Lower bound on N)**.** *For every sufficiently large integer N, the class $\mathfrak{G}_{N,8}$ of N-player games with eight actions per player and utilities in $[0,1]$ satisfies*

$$\inf_{\eta>0} \sup_{G \in \mathfrak{G}_{N,8}} \sup_{T \geq 1} \max_{i \in [N]} \mathrm{Reg}_i^G(T) \geq \frac{\log 8}{1000}\, 2^{\lfloor N/2 \rfloor} - \frac{3}{2}.$$

The action sets are fixed before η is chosen; players' utility functions may depend on η.

The proof is deferred to Section B.1. This lower bound also extends to general OFTRL under suitable assumptions on the regularizer; see Proposition B.2 for details.

3.2 DEPENDENCE ON THE NUMBER OF ACTIONS

Theorem 3.2 (Lower bound on $m_{\max}$)**.** *For every sufficiently large integer m, the class $\mathfrak{G}_{6,m}$ of six-player games with m actions per player and utilities in $[0,1]$ satisfies*

$$\inf_{\eta>0} \sup_{G \in \mathfrak{G}_{6,m}} \sup_{T \geq 1} \max_{i \in [6]} \mathrm{Reg}_i^G(T) \geq \frac{2}{\pi} \frac{2^{\lfloor m/4 \rfloor} \log m}{m^3} - 1.$$

The proof is deferred to Section B.2. Theorems 3.1 and 3.2 motivate ECHO-OFTRL, whose regret bound is linear in N and logarithmic in $m_{\max}$.

4 ECHO-OFTRL

For every player $i \in [N]$, set

$$p_i \coloneqq \max\{2, \lceil \log m_i \rceil\}, \qquad M_i \coloneqq m_i^{1/p_i}, \qquad r_i \coloneqq \frac{(p_i+1)M_i}{2}.$$

For $\boldsymbol{y} \in \mathbb{R}_{\geq 0}^{m_i}$, define the regularizer $\psi_i \colon \left\{\boldsymbol{y} \in \mathbb{R}_{\geq 0}^{m_i} \colon \|\boldsymbol{y}\|_{p_i/(p_i-1)} \leq \frac{1}{M_i}\right\} \to \mathbb{R}$,

$$\psi_i(\boldsymbol{y}) \coloneqq \min_{\substack{\boldsymbol{z} \geq 0 \\ \|\boldsymbol{y}+\boldsymbol{z}\|_{p_i/(p_i-1)} \leq 1/M_i}} \left\{ -2r_i \sum_{a_i=1}^{m_i} \sqrt{y(a_i) z(a_i)} \right\},$$

with value $+\infty$ outside its feasible domain. For every $\boldsymbol{\theta} \in \mathbb{R}^{m_i}$, define the regularized response

$$\boldsymbol{y}_i(\boldsymbol{\theta}) \coloneqq \operatorname*{argmax}_{\boldsymbol{y} \in \operatorname{dom} \psi_i} \{\langle \boldsymbol{\theta}, \boldsymbol{y} \rangle - \psi_i(\boldsymbol{y})\}, \tag{4.1}$$

and set

$$\lambda_i(\boldsymbol{\theta}) \coloneqq \sum_{a_i=1}^{m_i} y_i(\boldsymbol{\theta})(a_i), \qquad Q_i(\boldsymbol{\theta}) \coloneqq (\lambda_i(\boldsymbol{\theta}), \boldsymbol{y}_i(\boldsymbol{\theta})), \qquad \boldsymbol{x}_i(\boldsymbol{\theta}) \coloneqq \boldsymbol{y}_i(\boldsymbol{\theta})/\lambda_i(\boldsymbol{\theta}).$$

Lemma 4.1. *For every $i \in [N]$, $1 \le M_i \le e$ and*

$$\operatorname{dom} \psi_i = \left\{ \boldsymbol{y} \in \mathbb{R}^{m_i}_{\ge 0} : \|\boldsymbol{y}\|_{p_i/(p_i-1)} \le 1/M_i \right\}.$$

The range of ψ_i is r_i. For every $\boldsymbol{\theta} \in \mathbb{R}^{m_i}$, the maximizer $\boldsymbol{y}_i(\boldsymbol{\theta})$ in (4.1) is unique and belongs to the interior of $\operatorname{dom} \psi_i$. Moreover, $0 < \lambda_i(\boldsymbol{\theta}) < 1$ and $\nabla \psi_i(\boldsymbol{y}_i(\boldsymbol{\theta})) = \boldsymbol{\theta}$. In addition, $\boldsymbol{y}_i(\mathbf{0}) = \mathbf{1}/(2m_i)$. Defining $\operatorname{asinh}(z) \coloneqq \log\left(z + \sqrt{1+z^2}\right)$, we have, for every $a_i \in [m_i]$,

$$x_i(\boldsymbol{\theta})(a_i) = \frac{\exp\left(p_i \operatorname{asinh}(\theta(a_i)/(2r_i)) - \frac{1}{2}\log(1 + (\theta(a_i)/(2r_i))^2)\right)}{\sum_{b_i=1}^{m_i} \exp\left(p_i \operatorname{asinh}(\theta(b_i)/(2r_i)) - \frac{1}{2}\log(1 + (\theta(b_i)/(2r_i))^2)\right)}. \tag{4.2}$$

The proof is deferred to Section C. In particular, (4.2) gives a closed-form update rule of (4.1).

4.1 Lifted regret

For every player $i \in [N]$, define the centered payoff map by

$$\mathcal{G}_i(\boldsymbol{x}) \coloneqq \mathcal{U}_i(\boldsymbol{x}_{-i}) - \langle \mathcal{U}_i(\boldsymbol{x}_{-i}), \boldsymbol{x}_i \rangle \mathbf{1}.$$

For every timestep $t \ge 1$, the centered payoff is

$$\boldsymbol{g}_i^{(t)} \coloneqq \mathcal{G}_i(\boldsymbol{x}^{(t)}) = \boldsymbol{u}_i^{(t)} - \left\langle \boldsymbol{u}_i^{(t)}, \boldsymbol{x}_i^{(t)} \right\rangle \mathbf{1}.$$

Set $\boldsymbol{g}_i^{(t)} = \mathbf{0}$ for $t \le 0$ and write $\boldsymbol{g}_i = (\boldsymbol{g}_i^{(t)})_{t \in \mathbb{Z}}$. For every $t \ge 1$, $\left\|\boldsymbol{g}_i^{(t)}\right\|_\infty \le 1$ and $\left\langle \boldsymbol{g}_i^{(t)}, \boldsymbol{x}_i^{(t)} \right\rangle = 0$. The lifted strategy is $\boldsymbol{y}_i^{(t)} = \lambda_i^{(t)} \boldsymbol{x}_i^{(t)}$. Since $\mathbf{0}$ and every pure strategy scaled by $1/M_i$ belong to $\operatorname{dom} \psi_i$, every horizon $T \ge 1$ satisfies

$$[\operatorname{Reg}_i(T)]_+ = \max_{\substack{\widehat{\lambda}_i \in [0,1] \\ \widehat{\boldsymbol{x}}_i \in \Delta^{m_i}}} \widehat{\lambda}_i \sum_{t=1}^{T} \left\langle \boldsymbol{g}_i^{(t)}, \widehat{\boldsymbol{x}}_i - \boldsymbol{x}_i^{(t)} \right\rangle \le M_i \max_{\widehat{\boldsymbol{y}}_i \in \operatorname{dom} \psi_i} \sum_{t=1}^{T} \left\langle \boldsymbol{g}_i^{(t)}, \widehat{\boldsymbol{y}}_i - \boldsymbol{y}_i^{(t)} \right\rangle. \tag{4.3}$$

Thus, an optimistic FTRL bound on the lifted domain controls the positive individual regret. For $\boldsymbol{y}, \boldsymbol{y}'$ in the interior of the lifted domain, define the Bregman divergence as $D_{\psi_i}(\boldsymbol{y}', \boldsymbol{y}) \coloneqq \psi_i(\boldsymbol{y}') - \psi_i(\boldsymbol{y}) - \langle \nabla \psi_i(\boldsymbol{y}), \boldsymbol{y}' - \boldsymbol{y} \rangle$.

4.2 Prediction and update

We call the algorithm below *ECHO-OFTRL*, where ECHO stands for *EMA cascade for high-order optimism* and EMA denotes *exponential moving average*.

Fix a common learning rate $\eta > 0$. On bi-infinite sequences in a finite-dimensional normed space, let I be the identity operator and define the backward shift L by $(\mathsf{L}\boldsymbol{z})^{(t)} \coloneqq \boldsymbol{z}^{(t-1)}$ for every sequence $\boldsymbol{z}$ and integer t. Set $\mathsf{D} \coloneqq I - \mathsf{L}$. On bounded sequences, define

$$\rho \coloneqq 1 - \frac{1}{N}, \qquad \mathsf{A} \coloneqq (I - \rho\mathsf{L})^{-1} = \sum_{k=0}^{\infty} \rho^k \mathsf{L}^k, \qquad \delta \coloneqq \mathsf{AD} = \frac{I - \mathsf{L}}{I - \rho\mathsf{L}}. \tag{4.4}$$

For $z \in \mathbb{C}$ with $|z| < 1/\rho$, write $\delta(z) \coloneqq (1-z)/(1-\rho z)$ for the transfer function of δ. Since A and D commute, $\delta = \mathsf{DA}$. Thus, δ takes the one-step difference of the exponentially weighted sequence generated by A.

Algorithm 1 ECHO-OFTRL for player i

Require: Common learning rate $\eta > 0$ and $\rho = 1 - 1/N$

Initialize $\boldsymbol{g}_i^{(0)} \leftarrow \mathbf{0}$ and $\mathbf{ema}_{i,h}^{(1)} \leftarrow \mathbf{0}$ for every $h \in [N-1]$

for $t = 1, 2, \ldots$ **do**

  Form $\widehat{\boldsymbol{g}}_i^{(t)}$ using (4.6)

  Set $\boldsymbol{\theta}_i^{(t)} \leftarrow \eta \left(\sum_{s=1}^{t-1} \boldsymbol{g}_i^{(s)} + \widehat{\boldsymbol{g}}_i^{(t)} \right)$ and play $\boldsymbol{x}_i^{(t)}$ using (4.2) and (4.7)

  Observe $\boldsymbol{u}_i^{(t)}$ and compute $\boldsymbol{g}_i^{(t)} \leftarrow \boldsymbol{u}_i^{(t)} - \left\langle \boldsymbol{u}_i^{(t)}, \boldsymbol{x}_i^{(t)} \right\rangle \mathbf{1}$

  for $h = 1, \ldots, N-1$ **do**

    Update $\mathbf{ema}_{i,h}^{(t+1)}$ using (4.8) with running prefix sums

For every bounded sequence $\boldsymbol{z}$ that vanishes at nonpositive times and every integer $t \geq 1$,

$$((I - \delta)\boldsymbol{z})^{(t)} = (1 - \rho) \sum_{k=1}^{\infty} \rho^{k-1} \boldsymbol{z}^{(t-k)}.$$

For each $i \in [N]$, define the prediction sequence $\widehat{\boldsymbol{g}}_i$ and the corresponding prediction-error sequence $\boldsymbol{e}_i$ by

$$\widehat{\boldsymbol{g}}_i \coloneqq (I - \mathsf{D}\delta^{N-1})\boldsymbol{g}_i, \qquad \boldsymbol{e}_i \coloneqq \boldsymbol{g}_i - \widehat{\boldsymbol{g}}_i = \mathsf{D}\delta^{N-1}\boldsymbol{g}_i. \tag{4.5}$$

Thus, for every $i \in [N]$, the prediction error $\boldsymbol{e}_i$ is the one-step difference of $\delta^{N-1}\boldsymbol{g}_i$. Moreover,

$$I - \mathsf{D}\delta^{N-1} = \mathsf{L} + \mathsf{D}(I - \delta^{N-1}) = \mathsf{L} + \mathsf{D}(I - \delta) \sum_{h=1}^{N-1} \delta^{h-1}.$$

Together with the preceding formula for $I - \delta$, this identity shows that $\widehat{\boldsymbol{g}}_i^{(t)}$ depends only on feedback from timesteps strictly before t. The following EMA cascade computes this predictor recursively.

Every player $i \in [N]$ maintains $N - 1$ sequences $(\mathbf{ema}_{i,h}^{(t)})_{t \geq 1}$ in $\mathbb{R}^{m_i}$, indexed by $h \in [N-1]$. These sequences are defined by $\mathbf{ema}_{i,h}^{(t)} \coloneqq \left((I - \delta)\delta^{h-1}\boldsymbol{g}_i\right)^{(t)}$. Since $\boldsymbol{g}_i^{(t)} = \mathbf{0}$ for every $t \leq 0$, this definition gives $\mathbf{ema}_{i,h}^{(0)} = \mathbf{0}$ for every $i \in [N]$ and $h \in [N-1]$. For every $i \in [N]$, $h \in [N-1]$, and $t \geq 1$,

$$\boldsymbol{g}_i^{(t)} - \sum_{\ell=1}^{h-1} \mathbf{ema}_{i,\ell}^{(t)} = \left(\left(I - (I - \delta) \sum_{\ell=1}^{h-1} \delta^{\ell-1} \right) \boldsymbol{g}_i \right)^{(t)} = \left(\delta^{h-1}\boldsymbol{g}_i\right)^{(t)},$$

where the last equality follows from the geometric series identity. In particular, $\sum_{h=1}^{N-1} \mathbf{ema}_{i,h} = (I - \delta^{N-1})\boldsymbol{g}_i$. Hence, at every timestep $t \geq 1$, player i forms the predictor and plays according to

$$\widehat{\boldsymbol{g}}_i^{(t)} = \boldsymbol{g}_i^{(t-1)} + \sum_{h=1}^{N-1} \mathbf{ema}_{i,h}^{(t)} - \sum_{h=1}^{N-1} \mathbf{ema}_{i,h}^{(t-1)}, \tag{4.6}$$

$$\boldsymbol{\mu}_i^{(t)} = \left(\lambda_i^{(t)}, \lambda_i^{(t)} \boldsymbol{x}_i^{(t)} \right) \coloneqq Q_i \left(\eta \left(\sum_{s=1}^{t-1} \boldsymbol{g}_i^{(s)} + \widehat{\boldsymbol{g}}_i^{(t)} \right) \right). \tag{4.7}$$

By (4.4), $(I - \rho\mathsf{L})(I - \delta) = (1 - \rho)\mathsf{L}$. Consequently, for every bounded bi-infinite sequence $\boldsymbol{z}$ in a finite dimensional normed space and every integer $t \geq 0$,

$$((I - \delta)\boldsymbol{z})^{(t+1)} = \rho\left((I - \delta)\boldsymbol{z}\right)^{(t)} + (1 - \rho)\boldsymbol{z}^{(t)}.$$

Applying this identity with $\boldsymbol{z} = \delta^{h-1}\boldsymbol{g}_i$ and using the preceding relation yields the recursive cascade update below. Thus, after observing $\boldsymbol{u}_i^{(t)} = \mathcal{U}_i(\boldsymbol{x}_{-i}^{(t)})$ and computing $\boldsymbol{g}_i^{(t)}$, player i updates all $N - 1$ EMA vectors according to

$$\mathbf{ema}_{i,h}^{(t+1)} = \rho\mathbf{ema}_{i,h}^{(t)} + (1 - \rho) \left(\boldsymbol{g}_i^{(t)} - \sum_{\ell=1}^{h-1} \mathbf{ema}_{i,\ell}^{(t)} \right), \qquad h \in [N-1]. \tag{4.8}$$

For every $i \in [N]$, extend $\boldsymbol{x}_i^{(t)}$ and $\lambda_i^{(t)}$ to nonpositive timesteps by setting $\boldsymbol{x}_i^{(t)} := \boldsymbol{x}_i(\mathbf{0})$ and $\lambda_i^{(t)} := \lambda_i(\mathbf{0})$ for every $t \le 0$, and write $\boldsymbol{x}_i := (\boldsymbol{x}_i^{(t)})_{t\in\mathbb{Z}}$ and $\boldsymbol{\lambda}_i := (\lambda_i^{(t)})_{t\in\mathbb{Z}}$. By maintaining the prefix sums in (4.8), player i can implement the cascade using $O(Nm_i)$ memory and arithmetic per timestep. Moreover, Lemma 4.1 gives $0 < \lambda_i^{(t)} < 1$ for every $i \in [N]$ and $t \ge 1$, so $\boldsymbol{x}_i^{(t)}$ is well defined. The complete procedure is summarized in Algorithm 1.

4.3 Main result

Theorem 4.2. *In every N-player finite normal-form game with utilities in $[0,1]$, suppose every player runs Algorithm 1 with the common learning rate $\eta = 1/(2000N)$. Recall that $r_i = (p_i+1)m_i^{1/p_i}/2$ for every $i \in [N]$. Then, for every $i \in [N]$ and horizon $T \ge 1$,*

$$[\mathrm{Reg}_i(T)]_+ \le m_i^{1/p_i}\left(2000Nr_i + \sum_{j=1}^{N} r_j + \frac{3}{40N}\right) \le O\left(N \log m_{\max}\right).$$

A proof sketch appears in Section 5.3; the formal proof in Section D.2 specializes Theorem D.1 to ECHO-OFTRL. Adversarial regret guarantees for the same update with a regret-dependent learning rate are given in Section F.

5 ECHO-OFTRL with general filters

We extend ECHO-OFTRL by allowing an arbitrary analytic prediction-error filter, while keeping the response Q_i from Section 4.

5.1 Response and prediction

In this section, we consider filters satisfying the following assumption.

Assumption 5.1. *The prediction-error filter $F(z) = \sum_{s\ge 0} f_s z^s$, with $f_s \in \mathbb{R}$, satisfies the following conditions:*

1. ***Validity.*** *$f_0 = 1$. This ensures that the predictor $\widehat{\boldsymbol{g}}_i^{(t)} = ((I - F(\mathsf{L}))\boldsymbol{g}_i)^{(t)}$ depends only on past observations, as required in online learning.*
2. ***Boundedness.*** *There exists $R > 1$ such that the series defining $F(z)$ converges for every $z \in \mathbb{R}$ with $|z| \le R$. This ensures that the predictions are uniformly bounded whenever the payoff sequence is uniformly bounded.*

Let k be the largest integer in $\{0,\dots,N\}$ such that $\left.\frac{d^{h-1}F}{dz^{h-1}}\right|_{z=1} = 0$ for every integer $h \in [k]$. In particular, $k = 0$ if and only if $F(1) \ne 0$. Let $\mathsf{F} = F(\mathsf{L})$. Fix a common learning rate $\eta > 0$ and set $\boldsymbol{g}_i^{(t)} = \mathbf{0}$ for $t \le 0$. At each timestep $t \ge 1$, player i forms the prediction

$$\widehat{\boldsymbol{g}}_i^{(t)} := ((I - \mathsf{F})\boldsymbol{g}_i)^{(t)} = -\sum_{s=1}^{t-1} f_s \boldsymbol{g}_i^{(t-s)}. \tag{5.1}$$

Player i then uses this prediction in (4.7). The prediction error satisfies $\boldsymbol{e}_i := \boldsymbol{g}_i - \widehat{\boldsymbol{g}}_i = \mathsf{F}\boldsymbol{g}_i$.

5.2 Main result

Theorem 5.2 (Informal)**.** *Suppose Assumption 5.1 holds. If all players use* (4.7) *and* (5.1) *with the common filter F and a sufficiently small common learning rate $\eta > 0$, then every player $i \in [N]$ satisfies, simultaneously for every $T \ge 1$,*

$$[\mathrm{Reg}_i(T)]_+ \le O_{F,N}\left(\frac{\log m_{\max}}{\eta} + (N-k)\eta^{2k+1}T\right).$$

The learning-rate threshold and implicit constant depend only on F, N. In particular, if $k = N$, a learning rate independent of T gives individual regret at most $O_{F,N}(\log m_{\max})$.

The formal version and its proof appear in Section D.

For $k < N$, tuning the learning rate to a prescribed horizon yields a regret bound with horizon dependence $T^{1/(2k+2)}$, as illustrated in Figure 1. In particular, when $k = 1$, this recovers the $O(T^{1/4})$ rate obtained from the RVU analysis with vanilla optimism $\widehat{\boldsymbol{g}}_i^{(t)} = \boldsymbol{g}_i^{(t-1)}$ (Syrgkanis et al., 2015). The proof sketch below explains how the zero multiplicity determines the number of recursive steps.

5.3 Proof sketch

Fix $T \geq 1$. Extend scores, strategies, and masses constantly to nonpositive times using their values at zero score. For $i \in [N]$, define

$$\mathcal{P}_{i,T} \coloneqq \sum_{t=1}^{T} D_{\psi_i}(\boldsymbol{y}_i^{(t)}, \boldsymbol{y}_i^{(t-1)}), \qquad \mathcal{P}_T \coloneqq \sum_{i=1}^{N} \mathcal{P}_{i,T}.$$

Recall that $\mathsf{D} = I - \mathsf{L}$. For $m \geq 2$, positive $\boldsymbol{x} \in \Delta^m$, and $\boldsymbol{v} \in \mathbb{R}^m$, write $\|\boldsymbol{v}\|_{\boldsymbol{x}}^2 = \sum_a x(a)v(a)^2$ and $\|\boldsymbol{v}\|_{\boldsymbol{x}^{-1}}^2 = \sum_a v(a)^2/x(a)$. By (4.3) and (D.7), for sufficiently small $\eta > 0$,

$$\sum_{i=1}^{N} \frac{[\mathrm{Reg}_i(T)]_+}{M_i} \leq \sum_{i=1}^{N} \frac{r_i}{\eta} + 3\eta \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\lambda_i^{(t)}\right)^{3/2} \left\| (\mathsf{F}\boldsymbol{g}_i)^{(t)} \right\|_{\boldsymbol{x}_i^{(t)}}^2 - \frac{\mathcal{P}_T}{2\eta}.$$

For $k = 0$, bounded prediction errors give the claimed bound directly; see (D.24). Assume $k \geq 1$ below. Fix R from Assumption 5.1, choose $\tau \geq N$ depending only on F, N with $\tau > 1/(2(R^2-1))$, and set $\rho_\tau \coloneqq 1 - 1/\tau$. Define

$$F_h(z) \coloneqq \frac{F(z)(1-\rho_\tau z)^{k-h}}{(1-z)^{k-h}} \qquad (0 \leq h \leq k). \tag{5.2}$$

Let $\mathsf{F}_h \coloneqq F_h(\mathsf{L})$, so $\mathsf{F}_k = \mathsf{F}$. For $k \geq 1$ and $i \in [N]$, the recursion is

$$\mathsf{F}\boldsymbol{g}_i \longrightarrow (\mathsf{F}_k \boldsymbol{x}_j)_{j\in[N]} \longrightarrow (\mathsf{F}_{k-1}\boldsymbol{g}_j)_{j\in[N]} \longrightarrow \cdots \longrightarrow (\mathsf{F}_1 \boldsymbol{x}_j)_{j\in[N]}.$$

Arrows denote bounds on weighted squared norms, up to movement and initialization terms. For $1 \leq h \leq k$, the two steps are as follows.

$(\mathsf{F}_h\boldsymbol{g}_j)_{j\in[N]} \longrightarrow (\mathsf{F}_h\boldsymbol{x}_j)_{j\in[N]}$. Multilinearity of the centered payoff maps gives this step, with error at most $O_{F,N}(1 + \eta^2\mathcal{P}_T)$; see (D.14) and (D.18).

$(\mathsf{F}_h\boldsymbol{x}_j)_{j\in[N]} \longrightarrow (\mathsf{F}_{h-1}\boldsymbol{g}_j)_{j\in[N]}$. For $i \in [N]$, the update gives $\mathsf{D}\boldsymbol{\theta}_i = \eta(I - \mathsf{DF})\boldsymbol{g}_i$. Since $\mathsf{F}_h = (I - \rho_\tau\mathsf{L})^{-1}\mathsf{F}_{h-1}\mathsf{D}$,

$$\mathsf{F}_h\boldsymbol{x}_i \overset{(i)}{\longrightarrow} (I - \rho_\tau\mathsf{L})^{-1}\mathsf{F}_{h-1}\mathsf{D}\boldsymbol{\theta}_i \overset{(ii)}{=} \eta(I - \rho_\tau\mathsf{L})^{-1}(I - \mathsf{DF})\mathsf{F}_{h-1}\boldsymbol{g}_i.$$

By Lemma D.2, (i) contributes a factor proportional to $\sqrt{\lambda_i^{(t)}}$ in the local norm, while (ii) supplies η. For sufficiently small η, weighted convolution and (D.13) give, for $i \in [N]$ and $1 \leq h \leq k$,

$$\sum_{t=1}^{T} \sqrt{\lambda_i^{(t)}} \left\| (\mathsf{F}_h\boldsymbol{x}_i)^{(t)} \right\|_{(\boldsymbol{x}_i^{(t)})^{-1}}^2 \leq O_{F,N}(\mathcal{P}_{i,T}). \tag{5.3}$$

Starting at player $i \in [N]$, the weight $(\lambda_i^{(t)})^{3/2} \leq \sqrt{\lambda_i^{(t)}}$. Since each response step at player $j \in [N]$ gains $\lambda_j^{(t)}$, we retain a weight bounded by $\sqrt{\lambda_j^{(t)}}$ for every visited player j. Revisiting a player therefore invokes (5.3); each continuing step adds a new player and lowers h. Figure 2 (left) illustrates this recursion for $N = k = 3$.

At $h = 1$, k distinct players have been visited, since any earlier revisit is controlled by (5.3). If $k = N$, the pigeonhole principle controls every terminal term by (5.3). Otherwise, each of the $N - k$ unvisited players contributes at most $O_{F,N}(\eta^2 T)$ by (D.5). Finally,

$$\eta \sum_{i=1}^{N} \sum_{t=1}^{T} (\lambda_i^{(t)})^{3/2} \left\| (\mathsf{F}\boldsymbol{g}_i)^{(t)} \right\|_{\boldsymbol{x}_i^{(t)}}^2 \leq O_{F,N}\left(\eta + \eta\mathcal{P}_T + (N-k)\eta^{2k+1}T\right).$$

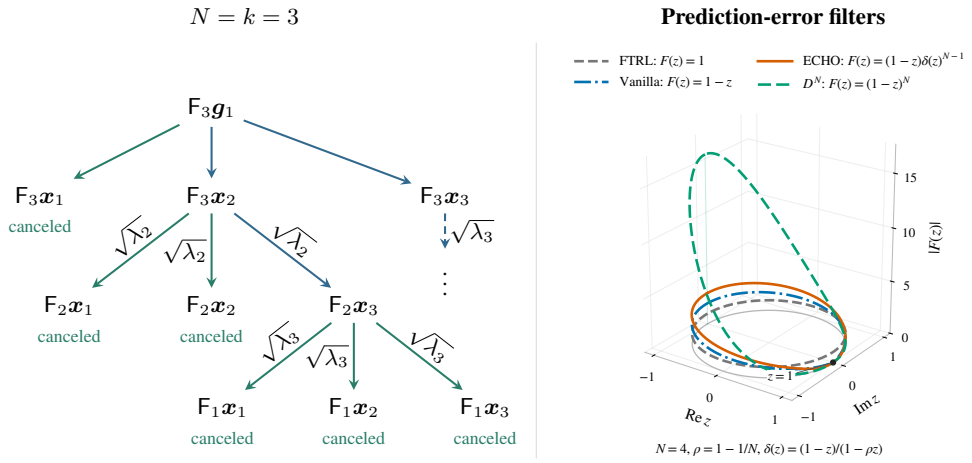

Figure 2: (Left) An illustration of how cancellation of $\mathsf{F}_h\boldsymbol{x}_i$ works in (5.3). (Right) Prediction-error filters on the unit circle for $N = 4$ and $\rho = 3/4$. Each curve plots $(\operatorname{Re} z, \operatorname{Im} z, |F(z)|)$ for $|z| = 1$. The unsmoothed difference D^N corresponds to $F(z) = (1-z)^N$ and reaches height $2^N = 16$ at $z = -1$, while ECHO reaches approximately 2.99.

Substituting into the regret bound, $-\mathcal{P}_T/(2\eta)$ absorbs the $O_{F,N}(\eta\mathcal{P}_T)$ term for small η. Since $M_i \le e$ and $r_i \le e(1 + \log m_i)$, Theorem 5.2 follows.

5.4 Examples

For every horizon $T \ge 1$, the following filters give the stated bounds for every player $i \in [N]$ when all players use the response Q_i from Section 4 with a suitable common learning rate.

- **FTRL.** $F(z) = 1$ and $k = 0$. $[\mathrm{Reg}_i(T)]_+ \le O\left(\sqrt{T \log m_{\max}}\right)$.
- **Vanilla OFTRL.** $F(z) = 1 - z$ and $k = 1$. $[\mathrm{Reg}_i(T)]_+ \le O\big(N \log m_{\max} + N^{1/4}(\log m_{\max})^{3/4} T^{1/4}\big)$.
- **ECHO-OFTRL (ours).** $F(z) = (1-z)\delta(z)^{N-1}$ with δ from (4.4) and $k = N$. $[\mathrm{Reg}_i(T)]_+ \le O\left(N \log m_{\max}\right)$.

The filter-norm estimates are proved in Section D.3 and Section D.2, respectively. Bounds for adversarial payoff sequences appear in Section D.4.

Figure 2 (right) compares these prediction-error filters with the unsmoothed difference D^N. For $\omega \in [-\pi, \pi]$, the height $\left|F(e^{i\omega})\right|$ is the amplitude multiplier when the filter acts on the scalar sequence $(e^{-i\omega t})_{t\in\mathbb{Z}}$.

6 Conclusion

We established horizon-independent individual regret with mild dependence on game size in finite normal-form games with utilities in $[0, 1]$ and full-information feedback. Standard Optimistic Hedge can incur regret exponential in either the number of players or the number of actions. ECHO-OFTRL combines an EMA cascade for high-order prediction with nonnegative lifted regret bounds to guarantee individual regret at most $O(N \log m_{\max})$ in self-play. Its updates are deterministic, uncoupled, and available in closed form.

More generally, our analysis relates the regret rate to the order of the prediction-error filter's zero at one. A regret-dependent learning rate also preserves the self-play guarantee while ensuring adversarial regret at most $O(N \log m_{\max} + \sqrt{T \log m_{\max}})$. Extending these guarantees beyond finite normal-form games and full-information feedback remains a direction for future work.

AI USE STATEMENT

We used ChatGPT 5.5, ChatGPT 5.6 Sol, and ChatGPT 6 Astra to explore research ideas and to assist in developing and writing initial versions of the proofs. We also used these tools to draft and polish the manuscript, assist in developing illustrative figures (*e.g.*, Figure 2), and help identify relevant literature to ensure that we discussed it thoroughly. The authors subsequently revised, verified, and substantially streamlined both the proofs and the manuscript. We take full responsibility for the final content of this work, including all AI-assisted text and mathematical claims.

REPRODUCIBILITY STATEMENT

All proofs are included in the paper's main text and appendix. Assumptions are listed explicitly (*e.g.*, Assumption 5.1).

REFERENCES

Omar Abbadi, Rida Laraki, and Panayotis Mertikopoulos. Constant regret in general games via higher-order optimism. *arXiv preprint arXiv:2609.04113*, 2026.

Ioannis Anagnostides, Gabriele Farina, Christian Kroer, Chung-Wei Lee, Haipeng Luo, and Tuomas Sandholm. Uncoupled learning dynamics with $O(\log T)$ swap regret in multiplayer games. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2022.

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

Noam Brown and Tuomas Sandholm. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science*, 359(6374):418–424, 2018.

Noam Brown and Tuomas Sandholm. Superhuman AI for multiplayer poker. *Science*, 365(6456): 885–890, 2019.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.

Xi Chen and Binghui Peng. Hedging in games: Faster convergence of external and swap regrets. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2020.

Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, and Shenghuo Zhu. Online optimization with gradual variations. In *Conference on Learning Theory (COLT)*, 2012.

Constantinos Daskalakis and Noah Golowich. Fast rates for nonparametric online learning: from realizability to learning in games. In *Annual ACM Symposium on Theory of Computing (STOC)*, 2022.

Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich. Near-optimal no-regret learning in general games. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2021.

Gabriele Farina, Christian Kroer, Noam Brown, and Tuomas Sandholm. Stable-predictive optimistic counterfactual regret minimization. In *International Conference on Machine Learning (ICML)*, 2019.

Gabriele Farina, Ioannis Anagnostides, Haipeng Luo, Chung-Wei Lee, Christian Kroer, and Tuomas Sandholm. Near-optimal no-regret learning dynamics for general convex games. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2022a.

Gabriele Farina, Chung-Wei Lee, Haipeng Luo, and Christian Kroer. Kernelized multiplicative weights for 0/1-polyhedral games: Bridging the gap between learning in extensive-form and normal-form games. In *International Conference on Machine Learning*, pp. 6337–6357. PMLR, 2022b.

Dylan J. Foster, Zhiyuan Li, Thodoris Lykouris, Karthik Sridharan, and Éva Tardos. Learning in games: Robustness of fast convergence. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2016.

Amy Greenwald and Amir Jafari. A general class of no-regret learning algorithms and game-theoretic equilibria. In *Learning Theory and Kernel Machines: 16th Annual Conference on Learning Theory and 7th Kernel Workshop (COLT/Kernel)*, 2003.

Junsoo Ha. A horizon-independent regret bound for Optimistic Hedge in general-sum games. *arXiv preprint arXiv:2609.22839*, 2026a.

Junsoo Ha. A logarithmic regret bound for Optimistic Hedge in general-sum games. *arXiv preprint arXiv:2609.19677*, 2026b.

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

Elad Hazan. Introduction to online convex optimization. *Foundations and Trends in Optimization*, 2 (3-4):157–325, 2016.

Tung Mai. Constant swap regret in general-sum games via two-scale higher-order optimism. *arXiv preprint arXiv:2609.16751*, 2026.

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 (ICLR)*, 2019.

Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisý, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. DeepStack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, 356(6337):508–513, 2017.

Hervé Moulin and J-P Vial. Strategically zero-sum games: the class of games whose completely mixed equilibria cannot be improved upon. *International Journal of Game Theory*, 7(3-4):201–221, 1978.

Georgios Piliouras, Lillian Ratliff, Ryann Sim, and Stratis Skoulakis. Fast convergence of optimistic gradient ascent in network zero-sum extensive form games. In *International Symposium on Algorithmic Game Theory (SAGT)*, 2022a.

Georgios Piliouras, Ryann Sim, and Stratis Skoulakis. Beyond time-average convergence: Near-optimal uncoupled online learning via clairvoyant multiplicative weights update. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2022b.

Sasha Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2013.

Ashkan Soleymani and Georgios Piliouras. Multiplicative optimism for constant regret in games. *arXiv preprint arXiv:2609.21976*, 2026.

Ashkan Soleymani, Georgios Piliouras, and Gabriele Farina. Cautious optimism: A meta-algorithm for near-constant regret in general games. In *ACM Conference on Economics and Computation (EC)*, 2025a.

Ashkan Soleymani, Georgios Piliouras, and Gabriele Farina. Faster rates for no-regret learning in general games via cautious optimism. In *Annual ACM Symposium on Theory of Computing (STOC)*, 2025b.

Vasilis Syrgkanis, Alekh Agarwal, Haipeng Luo, and Robert E. Schapire. Fast convergence of regularized learning in games. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2015.

Jun-Kun Wang and Jacob D. Abernethy. Acceleration through optimistic no-regret dynamics. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2018.

A CONCURRENT WORK

We discuss recent concurrent results for finite general sum games under full information feedback. To our knowledge, none of these concurrent works provides lower bounds for Optimistic Hedge.

High-order optimism. Abbadi et al. (2026); Mai (2026) develop high-order optimistic methods with regret bounds independent of the horizon. Abbadi et al. (2026) prove $O(N^3 \log^2 m_{\max})$ individual external regret. Mai (2026) prove individual swap regret at most $O\left(\sqrt{N}\, m_{\max} \log m_{\max} \log^{5/2}(N m_{\max})\right)$. By comparison, ECHO-OFTRL achieves $O(N \log m_{\max})$ individual external regret and admits a closed-form update by Theorem 4.2. Our analysis also treats a general class of prediction-error filters, whereas these works analyze specific high-order predictor constructions.

Optimistic Hedge. For a prescribed horizon $T \geq 1$, Ha (2026b) prove that Optimistic Hedge with uniform initialization and a suitable learning rate guarantees $O(\sqrt{N} \log m_i \log(T+2))$ external regret to every player $i \in [N]$. Ha (2026a) prove that Optimistic Hedge with uniform initialization and a sufficiently small constant learning rate achieves constant individual regret, with the dependence on the number of players and their action set sizes left implicit. Our lower bounds quantify this implicit dependence in the worst case. The worst-case regret is at least $2^{\Omega(N)}$ with eight actions per player, and at least $2^{\Omega(m)}$ in six-player games with m actions per player, by Theorems 3.1 and 3.2. In contrast, ECHO-OFTRL guarantees $O(N \log m_{\max})$ individual regret.

Multiplicative optimism. Soleymani & Piliouras (2026) obtain $O(\sqrt{N} \log m_{\max})$ individual external regret using first-order optimism and improve the dependence on N relative to the $O(N \log m_{\max})$ bound for ECHO-OFTRL. The methods are otherwise different. Our analysis treats a general class of prediction-error filters within ECHO-OFTRL, while Soleymani & Piliouras (2026) analyze a specific algorithm based on multiplicative optimism.

B PROOFS OF THE LOWER BOUNDS FOR OPTIMISTIC HEDGE

The update in (Optimistic Hedge) has the equivalent cumulative form

$$x_i^{(t)}(a_i) = \frac{\exp\left(\eta\left[\sum_{s<t} u_i^{(s)}(a_i) + u_i^{(t-1)}(a_i)\right]\right)}{\sum_{b_i=1}^{m_i} \exp\left(\eta\left[\sum_{s<t} u_i^{(s)}(b_i) + u_i^{(t-1)}(b_i)\right]\right)}, \qquad \boldsymbol{u}_i^{(0)} = 0. \tag{B.1}$$

We retain the common constant learning rate and initialization from Section 3. In the proofs, vector norms without a subscript are Euclidean, and $\|\cdot\|_{\mathrm{op}}$ is the induced matrix norm. Action labels starting at zero are a relabeling of the sets $[m_i]$.

B.1 DEPENDENCE ON THE NUMBER OF PLAYERS

Theorem 3.1 (Lower bound on N)**.** *For every sufficiently large integer N, the class $\mathfrak{G}_{N,8}$ of N-player games with eight actions per player and utilities in $[0,1]$ satisfies*

$$\inf_{\eta>0} \sup_{G\in\mathfrak{G}_{N,8}} \sup_{T\geq 1} \max_{i\in[N]} \mathrm{Reg}_i^G(T) \geq \frac{\log 8}{1000}\, 2^{\lfloor N/2 \rfloor} - \frac{3}{2}.$$

The action sets are fixed before η is chosen; players' utility functions may depend on η.

Proof. First consider $N = 4 + 2L$ for a positive integer L, and fix

$$\eta_L = \frac{1000}{2^{L+2}}, \qquad \theta = \frac{2\pi}{2^{L+2}}, \qquad c = \alpha = b = \frac{1}{16}.$$

We construct a game with positive asymptotic linear regret at rate η_L. All error constants may depend on the fixed integer L, but not on time or the amplitude ε chosen below.

For $a \in \{0, \dots, 7\}$ and $\vartheta \in \mathbb{R}$, define

$$\theta_a = 2\pi a/8, \qquad \boldsymbol{\zeta}_a = (\cos\theta_a, \sin\theta_a),$$
$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad R_\vartheta = \cos\vartheta\, I + \sin\vartheta\, J.$$

For $\boldsymbol{z} \in \mathbb{R}^2$, define

$$p_a(\boldsymbol{z}) = \frac{e^{\boldsymbol{z}\cdot\boldsymbol{\zeta}_a}}{\sum_{b=0}^{7} e^{\boldsymbol{z}\cdot\boldsymbol{\zeta}_b}}, \qquad \Phi_8(\boldsymbol{z}) = \sum_{a=0}^{7} p_a(\boldsymbol{z})\boldsymbol{\zeta}_a. \tag{B.2}$$

The trigonometric moments of these eight vectors give

$$\Phi_8(\boldsymbol{z}) = \frac{\boldsymbol{z}}{2} - \frac{\|\boldsymbol{z}\|^2 \boldsymbol{z}}{16} + O(\|\boldsymbol{z}\|^5), \qquad \boldsymbol{J}_{\Phi_8}(\boldsymbol{z}) = \frac{I}{2} + O(\|\boldsymbol{z}\|^2).$$

For example, their second moment is $I/2$, and their averaged fourth power $(\boldsymbol{z}\cdot\boldsymbol{\zeta}_a)^4$ is $3\|\boldsymbol{z}\|^4/8$. Expanding the log of the denominator in (B.2) and differentiating gives the displayed expansion. If the expected utility of action a is $1/2 + \boldsymbol{\zeta}_a \cdot \boldsymbol{\iota}^{(t)}$ for $\boldsymbol{\iota}^{(t)} \in \mathbb{R}^2$, then Optimistic Hedge stays in the family $p(\boldsymbol{z})$ and satisfies

$$\boldsymbol{z}^{(1)} = 0, \qquad \boldsymbol{\iota}^{(0)} = 0, \qquad \boldsymbol{z}^{(t+1)} = \boldsymbol{z}^{(t)} + \eta_L \left(2\boldsymbol{\iota}^{(t)} - \boldsymbol{\iota}^{(t-1)}\right). \tag{B.3}$$

Choose real parameters a_c, a_s and $d > 0$ as specified below. For the first four players, read player indices modulo four and set

$$\begin{aligned}\mathcal{U}_i(a_1, \dots, a_4) =& \frac{1}{2} + a_c \cos(\theta_{a_i} - \theta_{a_{i+1}}) + a_s \sin(\theta_{a_i} - \theta_{a_{i+1}}) \\ & - c\cos(\theta_{a_{i+1}} - \theta_{a_{i+2}})\cos(\theta_{a_i} - \theta_{a_{i+3}}) + d\cos\theta_{a_i}.\end{aligned} \tag{B.4}$$

These utilities are independent of the other players. Cyclic symmetry preserves a common strategy $p(\boldsymbol{z}_0^{(t)})$. Independence of the three opponents gives

$$\mathbb{E}\left[(\boldsymbol{\zeta}_{a_{i+1}} \cdot \boldsymbol{\zeta}_{a_{i+2}})\boldsymbol{\zeta}_{a_{i+3}}\right] = \|\Phi_8(\boldsymbol{z})\|^2 \Phi_8(\boldsymbol{z}).$$

Since $\boldsymbol{\zeta}_a \cdot J\boldsymbol{\zeta}_b = \sin(\theta_a - \theta_b)$, the common expected utility parameter at every $t \geq 1$ is

$$\boldsymbol{\iota}_0^{(t)} = \left(a_c I + a_s J - c\left\|\Phi_8(\boldsymbol{z}_0^{(t)})\right\|^2 I\right)\Phi_8(\boldsymbol{z}_0^{(t)}) + d(1,0).$$

Thus the common source score follows (B.3); its first two scores are 0 and $2\eta_L d(1,0)$.

Lemma B.1. *Let $\Phi_{8,*}$ be a C^5 map from a neighborhood of $0 \in \mathbb{R}^2$ to $\mathbb{R}^2$ such that, for some $c_3 \in \mathbb{R}$,*

$$\Phi_{8,*}(\boldsymbol{z}) = \frac{\boldsymbol{z}}{2} + c_3\|\boldsymbol{z}\|^2\boldsymbol{z} + O(\|\boldsymbol{z}\|^5). \tag{B.5}$$

For every sufficiently large L and every sufficiently small $\varepsilon > 0$, there are real a_c, a_s and $d > 0$ with $\sqrt{a_c^2 + a_s^2} + c + d < 1/4$ such that (B.3)*, initialized by $\boldsymbol{z}_0^{(1)} = 0$ and $\boldsymbol{\iota}_0^{(0)} = 0$, with $\boldsymbol{\iota}_0^{(t)} = \iota_{8,*}(\boldsymbol{z}_0^{(t)})$ for $t \geq 1$ and*

$$\iota_{8,*}(\boldsymbol{z}) = \left(a_c I + a_s J - c\|\Phi_{8,*}(\boldsymbol{z})\|^2 I\right)\Phi_{8,*}(\boldsymbol{z}) + d(1,0),$$

satisfies the following bounds for some integer $t_* \geq 1$ and real phases $(\vartheta^{(t)})_{t\geq t_*}$, with the last two relations holding for every $t \geq t_*$:

$$\begin{aligned}\sup_{t\geq 1}\left\|\boldsymbol{z}_0^{(t)}\right\| \leq& O(\varepsilon), \\ \boldsymbol{z}_0^{(t)} =& \varepsilon R_{\vartheta^{(t)}}(1,0) + O(\varepsilon^2) + o_t(1), \\ \vartheta^{(t+1)} - \vartheta^{(t)} =& \theta + O(\varepsilon^2) + o_t(1).\end{aligned} \tag{B.6}$$

The thresholds and implicit constants may depend on $\Phi_{8,*}$. For each fixed ε, the terms $o_t(1)$ tend to zero as $t \to \infty$.

Proof. For real μ near zero, define

$$\begin{aligned}\boldsymbol{R}_\mu =&(1+\mu)R_\theta, \qquad \boldsymbol{C}_\mu = \frac{2}{\eta_L}(\boldsymbol{R}_\mu - I)(2I - \boldsymbol{R}_\mu^{-1})^{-1},\\ \boldsymbol{R}_{\mathrm{f},\mu} =&\frac{\eta_L}{2}\boldsymbol{C}_\mu \boldsymbol{R}_\mu^{-1},\\ \boldsymbol{V}_{1,\mu} =&\left(I - \frac{\eta_L}{2}\boldsymbol{C}_\mu \boldsymbol{R}_\mu^{-2}\right)^{-1}, \qquad \boldsymbol{V}_{2,\mu} = -\frac{\eta_L}{2}\boldsymbol{V}_{1,\mu}\boldsymbol{C}_\mu \boldsymbol{R}_\mu^{-1}.\end{aligned}$$

These are real matrices of the form $uI + vJ$. Their definitions give

$$\begin{aligned}\boldsymbol{R}_\mu =&I + \eta_L \boldsymbol{C}_\mu - \frac{\eta_L}{2}\boldsymbol{C}_\mu \boldsymbol{R}_\mu^{-1},\\ \boldsymbol{V}_{1,\mu} + \boldsymbol{V}_{2,\mu}\boldsymbol{R}_\mu^{-1} =&I.\end{aligned}$$

The cubic coefficient and its radial part at $\mu = 0$ are

$$\begin{aligned}\boldsymbol{C}_{\mathrm{cub},\mu} =&\eta_L \boldsymbol{V}_{1,\mu}\left(c_3\boldsymbol{C}_\mu - \frac{c}{8}I\right)\left(2I - (1+\mu)^{-2}\boldsymbol{R}_\mu^{-1}\right),\\ \ell =&\frac{1}{2}\operatorname{tr}(R_{-\theta}\boldsymbol{C}_{\mathrm{cub},0}), \qquad \frac{\ell}{\eta_L} \longrightarrow -\frac{c}{8} \quad \text{as } L \to \infty.\end{aligned}$$

Indeed, $\theta = (2\pi/1000)\eta_L$, so $\boldsymbol{C}_0 \to (4\pi/1000)J$, $\boldsymbol{V}_{1,0} \to I$, and $\boldsymbol{R}_{\mathrm{f},0} \to 0$. Since $\operatorname{tr} J = 0$, the limiting radial coefficient is $\frac{1}{2}\operatorname{tr}(c_3(4\pi/1000)J - (c/8)I) = -c/8$. Fix L large enough that $\ell < 0$, $\|\boldsymbol{C}_0\|_{\mathrm{op}} + c < 1/8$, and $\|\boldsymbol{R}_{\mathrm{f},0}\|_{\mathrm{op}} < 1/8$. Set $\mu = -\ell\varepsilon^2$, $d = \varepsilon^4$, and choose $a_c I + a_s J = \boldsymbol{C}_\mu$. For sufficiently small ε, $\|\boldsymbol{R}_{\mathrm{f},\mu}\|_{\mathrm{op}} \le 1/4$ and $\sqrt{a_c^2 + a_s^2} + c + d < 1/4$.

The expansion of $\Phi_{8,*}$ gives

$$\iota_{8,*}(\boldsymbol{z}) = \frac{1}{2}\boldsymbol{C}_\mu \boldsymbol{z} + \left(c_3\boldsymbol{C}_\mu - \frac{c}{8}I\right)\boldsymbol{z}\,\|\boldsymbol{z}\|^2 + d(1,0)^{\mathsf{T}} + O(\|\boldsymbol{z}\|^5).$$

The map $\boldsymbol{z} \mapsto \boldsymbol{z} - 2\boldsymbol{C}_\mu^{-1}\iota_{8,*}(\boldsymbol{z})$ is a contraction on the ball of radius $4\left\|\boldsymbol{C}_\mu^{-1}\right\|_{\mathrm{op}} d$ about zero for small d. Its fixed point satisfies $\iota_{8,*}(\boldsymbol{z}_*) = 0$ and $\boldsymbol{z}_* = -2d\boldsymbol{C}_\mu^{-1}(1,0)^{\mathsf{T}} + O(d^3)$. Define the centered remainder

$$\begin{aligned}\widehat{\iota}_{8,*}(\boldsymbol{z}) =&\iota_{8,*}(\boldsymbol{z}_* + \boldsymbol{z}) - \frac{1}{2}\boldsymbol{C}_\mu \boldsymbol{z}\\ =&\left(c_3\boldsymbol{C}_\mu - \frac{c}{8}I\right)\boldsymbol{z}\,\|\boldsymbol{z}\|^2 + O(d^2\,\|\boldsymbol{z}\| + d\,\|\boldsymbol{z}\|^2 + \|\boldsymbol{z}\|^5).\end{aligned}$$

For every $t \ge 2$, make the real linear change of coordinates

$$\begin{aligned}\boldsymbol{w}^{(t)} =&\boldsymbol{V}_{1,\mu}(\boldsymbol{z}_0^{(t)} - \boldsymbol{z}_*) + \boldsymbol{V}_{2,\mu}(\boldsymbol{z}_0^{(t-1)} - \boldsymbol{z}_*),\\ \boldsymbol{v}^{(t)} =&\boldsymbol{z}_0^{(t-1)} - \boldsymbol{z}_* - \boldsymbol{R}_\mu^{-1}(\boldsymbol{z}_0^{(t)} - \boldsymbol{z}_*).\end{aligned}$$

Its inverse is

$$\begin{aligned}\boldsymbol{z}_0^{(t)} - \boldsymbol{z}_* =&\boldsymbol{w}^{(t)} - \boldsymbol{V}_{2,\mu}\boldsymbol{v}^{(t)},\\ \boldsymbol{z}_0^{(t-1)} - \boldsymbol{z}_* =&\boldsymbol{R}_\mu^{-1}\boldsymbol{w}^{(t)} + (I - \boldsymbol{R}_\mu^{-1}\boldsymbol{V}_{2,\mu})\boldsymbol{v}^{(t)}.\end{aligned}$$

Writing $\boldsymbol{e}^{(t)} = \eta_L[2\widehat{\iota}_{8,*}(\boldsymbol{z}_0^{(t)} - \boldsymbol{z}_*) - \widehat{\iota}_{8,*}(\boldsymbol{z}_0^{(t-1)} - \boldsymbol{z}_*)]$, the two-step update gives the exact recurrences

$$\begin{aligned}\boldsymbol{w}^{(t+1)} =&\boldsymbol{R}_\mu \boldsymbol{w}^{(t)} + \boldsymbol{V}_{1,\mu}\boldsymbol{e}^{(t)},\\ \boldsymbol{v}^{(t+1)} =&\boldsymbol{R}_{\mathrm{f},\mu}\boldsymbol{v}^{(t)} - \boldsymbol{R}_\mu^{-1}\boldsymbol{e}^{(t)}.\end{aligned} \tag{B.7}$$

Let $r^{(t)} = \left\|\boldsymbol{w}^{(t)}\right\|$ and $s^{(t)} = \left\|\boldsymbol{v}^{(t)}\right\|$. The Taylor expansion above implies

$$\left\|\boldsymbol{e}^{(t)}\right\| \le O\left((r^{(t)} + s^{(t)})(d + r^{(t)} + s^{(t)})^2\right).$$

More precisely, substituting the inverse change of coordinates gives

$$\boldsymbol{w}^{(t+1)} =\boldsymbol{R}_\mu \boldsymbol{w}^{(t)} + \boldsymbol{C}_{\mathrm{cub},\mu}\boldsymbol{w}^{(t)}(r^{(t)})^2 + \boldsymbol{e}_{\mathrm{cub}}^{(t)},$$

$$\begin{aligned}\left\|\boldsymbol{e}_{\mathrm{cub}}^{(t)}\right\| \leq & O\left((r^{(t)}+s^{(t)})^2 s^{(t)}+(r^{(t)}+s^{(t)})^5\right) \\ & +O\left(d(r^{(t)}+s^{(t)})^2+d^2(r^{(t)}+s^{(t)})\right).\end{aligned} \tag{B.8}$$

The coefficient follows from $\left\|\boldsymbol{R}_\mu^{-1}\boldsymbol{w}\right\|^2=(1+\mu)^{-2}\|\boldsymbol{w}\|^2$; every other cubic term contains at least one factor $\boldsymbol{v}^{(t)}$.

The prescribed initial pair is $(\boldsymbol{z}_0^{(2)},\boldsymbol{z}_0^{(1)})=(2\eta_L d(1,0)^{\mathsf{T}},0)$. Its transformed first component is

$$\begin{aligned}\boldsymbol{w}^{(2)} =& \boldsymbol{V}_{1,\mu}\left(2\eta_L I+2\boldsymbol{C}_\mu^{-1}-\eta_L\boldsymbol{R}_\mu^{-1}\right)d(1,0)^{\mathsf{T}}+O(d^3) \\ =& 2\boldsymbol{V}_{1,\mu}\boldsymbol{R}_\mu\boldsymbol{C}_\mu^{-1}d(1,0)^{\mathsf{T}}+O(d^3).\end{aligned} \tag{B.9}$$

The leading matrix is invertible, whereas $s^{(2)}\leq O(d)$. Consequently there is a fixed $K_0\geq 1$ such that $0<r^{(2)}\leq O(d)$ and $s^{(2)}\leq K_0 r^{(2)}$ for small ε. By (B.7), whenever $s^{(t)}\leq K_0 r^{(t)}$, a fixed constant $C>0$ gives

$$\begin{aligned}\left|r^{(t+1)}-(1+\mu)r^{(t)}\right| \leq & C\left((r^{(t)})^2+d^2\right)r^{(t)}, \\ s^{(t+1)} \leq & \tfrac{1}{4}s^{(t)}+C\left((r^{(t)})^2+d^2\right)r^{(t)}.\end{aligned}$$

Choose a fixed $a\in(0,1/4)$ with $Ca^2\leq -\ell/4$. For small ε, $Cd^2\leq\mu/4$, so, while $r^{(t)}\leq a\varepsilon$ and $s^{(t)}\leq K_0 r^{(t)}$,

$$\begin{aligned}(1+\mu/2)r^{(t)} \leq & r^{(t+1)}\leq(1+2\mu)r^{(t)}, \\ s^{(t+1)} \leq & \tfrac{1}{4}K_0 r^{(t)}+\tfrac{\mu}{2}r^{(t)}\leq K_0 r^{(t+1)}.\end{aligned}$$

Thus (B.9) ensures a finite first time t_0 with $a\varepsilon\leq r^{(t_0)}\leq 2a\varepsilon$.

The second recurrence in (B.7) also controls $s^{(t)}$ independently of the ratio $s^{(t)}/r^{(t)}$. There is a fixed $C_1>0$ such that, for $r^{(t)}\leq 2\varepsilon$ and $s^{(t)}\leq\varepsilon$,

$$s^{(t+1)}\leq\tfrac{1}{4}s^{(t)}+C_1\varepsilon^3.$$

Set $K_1=2C_1+1$. Since $s^{(2)}\leq O(\varepsilon^4)$, induction gives $s^{(t)}\leq K_1\varepsilon^3$ as long as $r^{(t)}\leq 2\varepsilon$, because $K_1\varepsilon^3\leq\varepsilon$ and $K_1/4+C_1\leq K_1$ for small ε. In particular this bound holds at t_0.

For $a\varepsilon\leq r^{(t)}\leq 2\varepsilon$ and $s^{(t)}\leq K_1\varepsilon^3$, (B.8) has remainder $O(\varepsilon^5)$. Write $\boldsymbol{w}^{(t)}=r^{(t)}R_{\vartheta^{(t)}}(1,0)^{\mathsf{T}}$, choosing consecutive real phases with increments near θ. Since $\frac{1}{2}\operatorname{tr}(R_{-\theta}\boldsymbol{C}_{\mathrm{cub},\mu})=\ell+O(\mu)$, taking norms and angles gives

$$\begin{aligned}r^{(t+1)} =& r^{(t)}+\mu r^{(t)}+\ell(r^{(t)})^3+O(\varepsilon^5), \\ \vartheta^{(t+1)}-\vartheta^{(t)} =& \theta+O(\varepsilon^2).\end{aligned} \tag{B.10}$$

For $\widehat{r}^{(t)}=r^{(t)}/\varepsilon$, the radial recurrence becomes

$$\widehat{r}^{(t+1)}=\widehat{r}^{(t)}+\mu\widehat{r}^{(t)}(1-(\widehat{r}^{(t)})^2)+O(\mu\varepsilon^2).$$

For small ε, this implies $\widehat{r}^{(t+1)}\geq(1+\mu/4)\widehat{r}^{(t)}$ on $[a,4/5]$, and $\widehat{r}^{(t+1)}/\widehat{r}^{(t)}=1+O(\mu)$. Starting at t_0, the radius therefore enters $[4/5,6/5]$ in finite time. The scalar map $\widehat{r}\mapsto\widehat{r}+\mu\widehat{r}(1-\widehat{r}^2)$ is increasing there; its inward increments at the endpoints are $36\mu/125$ and $66\mu/125$. These dominate the $O(\mu\varepsilon^2)$ error. Hence this interval is preserved, together with $s^{(t)}\leq K_1\varepsilon^3$. On this interval,

$$\begin{aligned}\widehat{r}^{(t+1)}-1 =& \left[1-\mu\widehat{r}^{(t)}(\widehat{r}^{(t)}+1)\right](\widehat{r}^{(t)}-1)+O(\mu\varepsilon^2), \\ \left|\widehat{r}^{(t+1)}-1\right| \leq & (1-36\mu/25)\left|\widehat{r}^{(t)}-1\right|+O(\mu\varepsilon^2).\end{aligned}$$

Iteration gives $\limsup_t\left|r^{(t)}-\varepsilon\right|\leq O(\varepsilon^3)$. The inverse coordinate formula and $\|\boldsymbol{z}_*\|\leq O(d)$ give $\boldsymbol{z}_0^{(t)}=\boldsymbol{w}^{(t)}+O(\varepsilon^3)$, and all preceding iterates have norm $O(\varepsilon)$. Together with (B.10), these bounds prove (B.6). □

Apply Lemma B.1 with $\Phi_{8,*} = \Phi_8$ and $c_3 = -1/16$. Every utility in (B.4) lies in $[0, 1]$, because

$$\left| \mathcal{U}_i - \frac{1}{2} \right| \le \sqrt{a_c^2 + a_s^2} + c + d < \frac{1}{4}.$$

For each $j \in \{1, \dots, L\}$, add two eight-action players with indices $2j + 3$ and $2j + 4$. For either player $i \in \{2j + 3, 2j + 4\}$, let a' be its partner's action and let a_1, a_2 be the preceding pair's actions; for $j = 1$, use two of the source players. Set

$$\mathcal{U}_i(a, a', a_1, a_2) = \frac{1}{2} - \alpha \cos(\theta_a - \theta_{a'}) + b \cos(\theta_a - \theta_{a_1} - \theta_{a_2}). \tag{B.11}$$

The utilities lie in $[3/8, 5/8]$. The two players in each pair remain identical. Define the real polynomial map

$$\Upsilon(x, y) = (x^2 - y^2, 2xy), \qquad (x, y) \in \mathbb{R}^2.$$

The angle-addition formulas and independence of the two preceding players show that the common expected utility parameter is

$$\boldsymbol{\iota}_j^{(t)} = -\alpha \Phi_8(\boldsymbol{z}_j^{(t)}) + b\Upsilon(\Phi_8(\boldsymbol{z}_{j-1}^{(t)})). \tag{B.12}$$

If $\Phi_8(\boldsymbol{z}_{j-1}^{(t)}) = (x, y)$, the expected cosine and sine of the sum of the two preceding action angles are $x^2 - y^2$ and $2xy$. This uses two distinct opponents, so each pure payoff in (B.11) is a valid normal-form payoff. For every $\boldsymbol{z} \in \mathbb{R}^2$ and $\vartheta \in \mathbb{R}$,

$$\Upsilon(R_\vartheta \boldsymbol{z}) = R_{2\vartheta} \Upsilon(\boldsymbol{z}), \qquad \|\Upsilon(\boldsymbol{z})\| = \|\boldsymbol{z}\|^2. \tag{B.13}$$

Thus the forcing doubles the preceding angular frequency.

For $0 < \omega \le \pi/2$, define the real matrix

$$\begin{aligned} \boldsymbol{\Omega}(\omega) =& (R_\omega - I)(2I - R_{-\omega})^{-1} \\ =& \frac{(1 - \cos\omega)(2\cos\omega - 1)}{5 - 4\cos\omega} I + \frac{\sin\omega(3 - 2\cos\omega)}{5 - 4\cos\omega} J. \end{aligned} \tag{B.14}$$

The coefficient of J is strictly positive. Since $\det(uI + vJ) = u^2 + v^2$, the following vectors are well defined:

$$\boldsymbol{\nu}_0 = (1, 0), \qquad \boldsymbol{\nu}_j = \left(\frac{\alpha}{2} I + \frac{\boldsymbol{\Omega}(2^j \theta)}{\eta_L} \right)^{-1} \frac{b}{4} \Upsilon(\boldsymbol{\nu}_{j-1}), \quad 1 \le j \le L. \tag{B.15}$$

They are all nonzero by (B.13). We prove inductively that, for $1 \le j \le L$,

$$\begin{aligned} \sup_{t \ge 1} \left\| \boldsymbol{z}_j^{(t)} \right\| \le& O(\varepsilon^{2^j}), \\ \boldsymbol{z}_j^{(t)} =& \varepsilon^{2^j} R_{2^j \vartheta^{(t)}} \boldsymbol{\nu}_j + O(\varepsilon^{2^j + 1}) + o_t(1), \\ \boldsymbol{\iota}_j^{(t)} =& \frac{\boldsymbol{\Omega}(2^j \theta)}{\eta_L} \varepsilon^{2^j} R_{2^j \vartheta^{(t)}} \boldsymbol{\nu}_j + O(\varepsilon^{2^j + 1}) + o_t(1). \end{aligned} \tag{B.16}$$

For sufficiently large L, one has $\eta_L \alpha < 1$. The linear homogeneous part of the two-step update acts by

$$(\boldsymbol{z}, y) \longmapsto ((1 - \eta_L \alpha)\boldsymbol{z} + (\eta_L \alpha / 2) y, \boldsymbol{z}).$$

Set

$$q = \frac{1 - \eta_L \alpha + \sqrt{1 + (\eta_L \alpha)^2}}{2} \in (0, 1).$$

Its norm in $\|(\boldsymbol{z}, y)\|_* = \max\{\|\boldsymbol{z}\|, q\|y\|\}$ is q, because $(1 - \eta_L \alpha) + \eta_L \alpha / (2q) = q$. The derivative estimate for Φ_8 therefore gives Lipschitz constant at most $(1 + q)/2 < 1$ for the nonlinear homogeneous map on a sufficiently small ball. The preceding layer, if bounded by $O(\varepsilon^{2^{j-1}})$, supplies forcing

bounded by $O(\varepsilon^{2^j})$. The contraction leaves a ball of this order invariant, proving the first bound in (B.16) from the zero initial scores. At the first step, the preceding prediction is zero; it satisfies the same forcing bound.

For the remaining bounds, insert $\varepsilon^{2^j} R_{2^j\vartheta^{(t)}}\boldsymbol{\nu}_j$ into the recurrence. By (B.13) and the preceding layer's expansion,

$$\begin{aligned}\Upsilon(\Phi_8(\boldsymbol{z}_{j-1}^{(t)})) =& \frac{1}{4}\varepsilon^{2^j} R_{2^j\vartheta^{(t)}}\Upsilon(\boldsymbol{\nu}_{j-1}) + O(\varepsilon^{2^j+1}) + o_t(1),\\ R_{2^j\vartheta^{(t+1)}} =& R_{2^j\theta}R_{2^j\vartheta^{(t)}} + O(\varepsilon^2) + o_t(1).\end{aligned}$$

All matrices in (B.14) commute with rotations. The leading terms cancel because (B.15) gives

$$-\frac{\alpha}{2}\boldsymbol{\nu}_j + \frac{b}{4}\Upsilon(\boldsymbol{\nu}_{j-1}) = \frac{\boldsymbol{\Omega}(2^j\theta)}{\eta_L}\boldsymbol{\nu}_j,$$

and the definition of $\boldsymbol{\Omega}$ gives $(R_{2^j\theta} - I) = \boldsymbol{\Omega}(2^j\theta)(2I - R_{-2^j\theta})$. The nonlinear remainder from Φ_8 has order $\varepsilon^{3\cdot 2^j}$. Thus the residual has limsup at most $O(\varepsilon^{2^j+1})$. If $e^{(t)}$ is the distance in $\|\cdot\|_*$ between the actual and proposed two-step scores, the contraction implies

$$e^{(t+1)} \le \frac{1+q}{2}e^{(t)} + O(\varepsilon^{2^j+1}) + o_t(1).$$

It follows that $\limsup_t e^{(t)} \le O(\varepsilon^{2^j+1})$. Substituting into (B.12) proves the last bound in (B.16). The base case is (B.6).

For either player i_* in the last pair, the fixed uniform comparator earns $1/2$ at every timestep, since $\sum_a \boldsymbol{\zeta}_a = 0$. Its played utility is $1/2 + \Phi_8(\boldsymbol{z}_L^{(t)})\cdot\boldsymbol{\iota}_L^{(t)}$. Therefore

$$\mathrm{Reg}_{i_*}(T) \ge -\sum_{t=1}^{T}\Phi_8(\boldsymbol{z}_L^{(t)})\cdot\boldsymbol{\iota}_L^{(t)}.$$

Now $2^L\theta = \pi/2$, and the identities

$$\boldsymbol{\Omega}(\pi/2) = \frac{-I+3J}{5}, \qquad \boldsymbol{z}\cdot J\boldsymbol{z} = 0, \qquad \boldsymbol{z}\cdot\boldsymbol{\Omega}(\pi/2)\boldsymbol{z} = -\frac{\|\boldsymbol{z}\|^2}{5}$$

give, by (B.16),

$$\Phi_8(\boldsymbol{z}_L^{(t)})\cdot\boldsymbol{\iota}_L^{(t)} = -\frac{\|\boldsymbol{\nu}_L\|^2}{10\eta_L}\varepsilon^{2^{L+1}} + O(\varepsilon^{2^{L+1}+1}) + o_t(1).$$

Choose $\varepsilon > 0$ sufficiently small after fixing L. Since $\boldsymbol{\nu}_L \neq 0$, averaging over time yields

$$\liminf_{T\to\infty}\frac{\mathrm{Reg}_{i_*}(T)}{T} \ge \frac{\|\boldsymbol{\nu}_L\|^2}{20\eta_L}\varepsilon^{2^{L+1}} > 0.$$

The phase cancels pointwise; no periodicity or summability of phase errors is required.

For every $\eta \ge \eta_L$, replace each pure utility by $1/2 + (\eta_L/\eta)(\mathcal{U}_i - 1/2)$. The utilities remain in $[0,1]$, the strategy trajectory at rate η is unchanged, and regret is multiplied by $\eta_L/\eta > 0$. Hence its supremum over time is infinite. For $0 < \eta < \eta_L$, give one player one action with constant utility one and seven with constant utility zero. For the same action sets,

$$\begin{aligned}\sup_T \mathrm{Reg}_i(T) =& \frac{7}{8} + \sum_{t=2}^{\infty}\frac{7}{7+e^{\eta t}}\\ \ge& \frac{7}{8} + \int_2^{\infty}\frac{7}{7+e^{\eta s}}\,\mathrm{d}s \ge \frac{\log 8}{\eta} - \frac{9}{8} \ge \frac{2^{L+2}\log 8}{1000} - \frac{3}{2}.\end{aligned}$$

For general sufficiently large N, choose $L = \lfloor (N-4)/2 \rfloor$ and add at most one eight-action player with constant utilities. Then $2^{L+2} = 2^{\lfloor N/2 \rfloor}$, proving the theorem. □

B.1.1 Extension to general OFTRL

Let $\psi\colon \Delta^8 \to \mathbb{R}$ be continuous and strictly convex. In an N-player game with eight actions per player, define the regularized response for $\boldsymbol{\theta} \in \mathbb{R}^8$ and the OFTRL update for $i \in [N]$, with a common learning rate $\eta > 0$ and $\boldsymbol{u}_i^{(0)} = 0$, by

$$\begin{aligned} Q_\psi(\boldsymbol{\theta}) =& \operatorname*{argmax}_{\boldsymbol{x}\in\Delta^8} \{\langle \boldsymbol{\theta}, \boldsymbol{x}\rangle - \psi(\boldsymbol{x})\}, \\ \boldsymbol{x}_i^{(t)} =& Q_\psi\left(\eta\left[\sum_{s<t} \boldsymbol{u}_i^{(s)} + \boldsymbol{u}_i^{(t-1)}\right]\right) \qquad (t \geq 1). \end{aligned} \tag{B.17}$$

Proposition B.2. *Suppose ψ is invariant under permutations of the eight coordinates and is C^6 on a relative neighborhood of $\bar{\boldsymbol{x}} := \mathbf{1}/8$. Assume that, for some $\sigma_\psi > 0$, its Hessian on the tangent space satisfies*

$$\boldsymbol{v}^\top \nabla^2 \psi(\bar{\boldsymbol{x}})\boldsymbol{v} = \sigma_\psi \|\boldsymbol{v}\|^2 \qquad (\boldsymbol{v} \in \mathbb{R}^8,\ \langle \boldsymbol{v}, \mathbf{1}\rangle = 0).$$

Let $\widehat{\boldsymbol{x}}_0 = (1, 0, \ldots, 0) \in \Delta^8$, and write $\mathrm{Reg}_{i,\psi}^G(T)$ for regret under (B.17)*, with every player using this fixed ψ. For every sufficiently large integer N, depending only on ψ,*

$$\inf_{\eta>0} \sup_{G\in\mathfrak{G}_{N,8}} \sup_{T\geq 1} \max_{i\in[N]} \mathrm{Reg}_{i,\psi}^G(T) \geq \frac{\psi(\widehat{\boldsymbol{x}}_0) - \psi(\bar{\boldsymbol{x}})}{125\sigma_\psi} 2^{\lfloor N/2\rfloor} - 1.$$

Proof. Use the vectors $\boldsymbol{\zeta}_a$, rotations R_ϑ, and matrix J from Section B.1, and define, for $\boldsymbol{z} \in \mathbb{R}^2$,

$$\Phi_{8,\psi}(\boldsymbol{z}) = \sum_{a=0}^{7} Q_\psi\left(\frac{\sigma_\psi}{8}(\boldsymbol{\zeta}_b \cdot \boldsymbol{z})_{b=0}^7\right)(a)\boldsymbol{\zeta}_a.$$

Symmetry and strict convexity give $Q_\psi(0) = \bar{\boldsymbol{x}}$. The implicit function theorem applied to the first-order condition on $\mathbf{1}^\perp$ shows that Q_ψ is C^5 near zero, with derivative $\sigma_\psi^{-1} I$ on $\mathbf{1}^\perp$. Since $\sum_a \boldsymbol{\zeta}_a = 0$ and $\sum_a \boldsymbol{\zeta}_a\boldsymbol{\zeta}_a^\top = 4I$,

$$\Phi_{8,\psi}(0) = 0, \qquad \boldsymbol{J}_{\Phi_{8,\psi}}(0) = \frac{1}{8}\sum_{a=0}^{7} \boldsymbol{\zeta}_a\boldsymbol{\zeta}_a^\top = \frac{I}{2}.$$

Permuting the action labels gives $\Phi_{8,\psi}(R_{\pi/4}\boldsymbol{z}) = R_{\pi/4}\Phi_{8,\psi}(\boldsymbol{z})$ and equivariance under the reflection $(x, y) \mapsto (x, -y)$. In particular, $\Phi_{8,\psi}$ is odd. Reflection and quarter-turn symmetry make the homogeneous cubic term have components $c_3x^3 + c_3'xy^2$ and $c_3y^3 + c_3'x^2y$ for real coefficients c_3, c_3'; evaluating the $\pi/4$-rotation identity at $(1, 0)$ gives $c_3' = c_3$. Thus $\Phi_{8,\psi}$ satisfies (B.5), and

$$\boldsymbol{J}_{\Phi_{8,\psi}}(\boldsymbol{z}) = \frac{I}{2} + O(\|\boldsymbol{z}\|^2).$$

For feedback vectors $\boldsymbol{\iota}^{(t)} \in \mathbb{R}^2$ at $t \geq 1$, set $\boldsymbol{\iota}^{(0)} = 0$. If each action $a \in \{0, \ldots, 7\}$ has expected utility $1/2 + \boldsymbol{\zeta}_a \cdot \boldsymbol{\iota}^{(t)}$, invariance of Q_ψ under addition of a constant vector gives

$$\begin{aligned} \boldsymbol{z}^{(t)} =& \frac{8\eta}{\sigma_\psi}\left[\sum_{s<t} \boldsymbol{\iota}^{(s)} + \boldsymbol{\iota}^{(t-1)}\right], \\ \boldsymbol{x}^{(t)} =& Q_\psi\left(\frac{\sigma_\psi}{8}(\boldsymbol{\zeta}_a \cdot \boldsymbol{z}^{(t)})_{a=0}^7\right), \\ \boldsymbol{z}^{(t+1)} =& \boldsymbol{z}^{(t)} + \frac{8\eta}{\sigma_\psi}\left(2\boldsymbol{\iota}^{(t)} - \boldsymbol{\iota}^{(t-1)}\right), \qquad \boldsymbol{z}^{(1)} = 0. \end{aligned}$$

For a sufficiently large integer L, set $N = 4 + 2L$ and take

$$\eta_{\psi,L} = \frac{125\sigma_\psi}{2^{L+2}}, \qquad \eta_L = \frac{8\eta_{\psi,L}}{\sigma_\psi} = \frac{1000}{2^{L+2}}, \qquad \theta = \frac{2\pi}{2^{L+2}}.$$

Use the pure utilities in (B.4) and (B.11), with $c = \alpha = b = 1/16$. At rate $\eta_{\psi,L}$, the score recurrence is (B.3). For the first four players, Lemma B.1 applies with $\Phi_{8,*} = \Phi_{8,\psi}$, producing (B.6) from the prescribed initialization while keeping all pure utilities in $[0, 1]$.

For each subsequent pair $j \in \{1, \dots, L\}$, independence gives

$$\boldsymbol{\iota}_j^{(t)} = -\alpha\Phi_{8,\psi}(\boldsymbol{z}_j^{(t)}) + b\Upsilon(\Phi_{8,\psi}(\boldsymbol{z}_{j-1}^{(t)})).$$

The linear companion map and its norm $q < 1$ are the same as in the proof of (B.16), since $\boldsymbol{J}_{\Phi_{8,\psi}}(0) = I/2$. On a sufficiently small ball the nonlinear homogeneous map therefore has Lipschitz constant at most $(1 + q)/2$. With $\boldsymbol{\nu}_j$ from (B.15), the leading forcing and cancellation are

$$\begin{aligned}\Upsilon(\Phi_{8,\psi}(\boldsymbol{z}_{j-1}^{(t)})) =& \frac{1}{4}\varepsilon^{2^j} R_{2^j\vartheta^{(t)}}\Upsilon(\boldsymbol{\nu}_{j-1}) + O(\varepsilon^{2^j+1}) + o_t(1),\\ -\frac{\alpha}{2}\boldsymbol{\nu}_j + \frac{b}{4}\Upsilon(\boldsymbol{\nu}_{j-1}) =& \frac{\boldsymbol{\Omega}(2^j\theta)}{\eta_L}\boldsymbol{\nu}_j.\end{aligned}$$

The response remainder is $O(\|\boldsymbol{z}\|^3)$, so the geometric recurrence for the error in that proof again gives (B.16). In particular, for either player i_* in the last pair,

$$\Phi_{8,\psi}(\boldsymbol{z}_L^{(t)}) \cdot \boldsymbol{\iota}_L^{(t)} = -\frac{\|\boldsymbol{\nu}_L\|^2}{10\eta_L}\varepsilon^{2^{L+1}} + O(\varepsilon^{2^{L+1}+1}) + o_t(1).$$

The uniform comparator earns $1/2$ each timestep. Choosing $\varepsilon > 0$ sufficiently small after ψ and L therefore gives

$$\liminf_{T\to\infty} \frac{\mathrm{Reg}_{i_*,\psi}^G(T)}{T} \geq \frac{\|\boldsymbol{\nu}_L\|^2}{20\eta_L}\varepsilon^{2^{L+1}} > 0.$$

For every $\eta \geq \eta_{\psi,L}$, replacing each utility $\mathcal{U}_i$ by $1/2 + (\eta_{\psi,L}/\eta)(\mathcal{U}_i - 1/2)$ preserves the trajectory and multiplies regret by $\eta_{\psi,L}/\eta > 0$. Hence the supremum over time is infinite for all such rates.

For $0 < \eta < \eta_{\psi,L}$, give one player constant utility one on action zero and zero on the other seven actions, and give every other player constant utilities. For $s \geq 0$, set $r(s) = 1 - Q_\psi(s\widehat{\boldsymbol{x}}_0)(0)$, and define $\psi^*(\boldsymbol{\theta}) = \max_{\boldsymbol{x}\in\Delta^8}\{\langle\boldsymbol{\theta}, \boldsymbol{x}\rangle - \psi(\boldsymbol{x})\}$. Uniqueness of the maximizer gives $\mathrm{d}\psi^*(s\widehat{\boldsymbol{x}}_0)/\mathrm{d}s = Q_\psi(s\widehat{\boldsymbol{x}}_0)(0)$, so convexity makes r nonincreasing with values in $[0, 1]$. For every $S \geq 0$,

$$\begin{aligned}\int_0^S r(s)\,\mathrm{d}s =& S - \psi^*(S\widehat{\boldsymbol{x}}_0) - \psi(\bar{\boldsymbol{x}}),\\ S - \psi^*(S\widehat{\boldsymbol{x}}_0) =& \min_{\boldsymbol{x}\in\Delta^8}\{S(1 - x(0)) + \psi(\boldsymbol{x})\}.\end{aligned}$$

Comparison with $\widehat{\boldsymbol{x}}_0$ forces every minimizer to converge to $\widehat{\boldsymbol{x}}_0$ as $S \to \infty$. The minimum lies between ψ evaluated at its minimizer and $\psi(\widehat{\boldsymbol{x}}_0)$, so continuity gives $\int_0^\infty r(s)\,\mathrm{d}s = \psi(\widehat{\boldsymbol{x}}_0) - \psi(\bar{\boldsymbol{x}}) > 0$. The OFTRL trajectory gives, for every $T \geq 1$,

$$\mathrm{Reg}_{1,\psi}^G(T) = r(0) + \sum_{t=2}^{T} r(\eta t) \geq \frac{1}{\eta}\int_0^{\eta(T+1)} r(s)\,\mathrm{d}s - r(\eta).$$

Taking the supremum over T yields

$$\sup_T \mathrm{Reg}_{1,\psi}^G(T) \geq \frac{\psi(\widehat{\boldsymbol{x}}_0) - \psi(\bar{\boldsymbol{x}})}{\eta} - 1 > \frac{\psi(\widehat{\boldsymbol{x}}_0) - \psi(\bar{\boldsymbol{x}})}{125\sigma_\psi}2^{L+2} - 1.$$

For general sufficiently large N, take $L = \lfloor (N - 4)/2 \rfloor$ and add at most one player with constant utilities. Then $2^{L+2} = 2^{\lfloor N/2 \rfloor}$. □

B.2 Dependence on the number of actions

Theorem 3.2 (Lower bound on $m_{\max}$)**.** *For every sufficiently large integer m, the class $\mathfrak{G}_{6,m}$ of six-player games with m actions per player and utilities in $[0, 1]$ satisfies*

$$\inf_{\eta>0}\sup_{G\in\mathfrak{G}_{6,m}}\sup_{T\geq 1}\max_{i\in[6]}\mathrm{Reg}_i^G(T) \geq \frac{2}{\pi}\frac{2^{\lfloor m/4 \rfloor}\log m}{m^3} - 1.$$

Proof. Set $L = \lfloor m/4 \rfloor - 1$. Partition $4(L+1)$ actions into blocks indexed by $j = 0, \dots, L$, each containing four actions. For each block j, let $\boldsymbol{\zeta}_{j,a} \in \mathbb{R}^2$ be respectively $(1,0), (-1,0), (0,1), (0,-1)$ on its four actions, and zero on all other actions. For $\boldsymbol{z} = (\boldsymbol{z}_0, \dots, \boldsymbol{z}_L) \in (\mathbb{R}^2)^{L+1}$, define

$$(\boldsymbol{V}\boldsymbol{z})_a = \sum_{j=0}^{L} \boldsymbol{\zeta}_{j,a} \cdot \boldsymbol{z}_j, \qquad p_a(\boldsymbol{z}) = \frac{\exp\left(m(\boldsymbol{V}\boldsymbol{z})_a/2\right)}{\sum_{b=1}^{m} \exp\left(m(\boldsymbol{V}\boldsymbol{z})_b/2\right)},$$
$$\Phi_{m,j}(\boldsymbol{z}) = \sum_{a=1}^{m} p_a(\boldsymbol{z}) \boldsymbol{\zeta}_{j,a}, \qquad \Phi_m(\boldsymbol{z}) = (\Phi_{m,0}(\boldsymbol{z}), \dots, \Phi_{m,L}(\boldsymbol{z})). \tag{B.18}$$

Thus $\boldsymbol{V}^\mathsf{T}\boldsymbol{V} = 2I$, $\boldsymbol{V}^\mathsf{T}\mathbf{1} = 0$, $\Phi_m(0) = 0$, and $\boldsymbol{J}_{\Phi_m}(0) = I$. All error constants below may depend on the fixed m, but not on time or the amplitude ε chosen below.

Swapping two opposite actions makes the corresponding component of Φ_m odd in that coordinate and all other components even. In particular, the quadratic Taylor term vanishes. Integrating each component in its own coordinate and expanding its derivative gives a constant $C_m > 0$ and a neighborhood of zero on which

$$\begin{aligned} \|\Phi_{m,j}(\boldsymbol{z}) - \boldsymbol{z}_j\| &\le C_m \|\boldsymbol{z}_j\| \|\boldsymbol{z}\|^2 \qquad (0 \le j \le L), \\ \|\Phi_{m,0}(\boldsymbol{z}) - \Phi_{m,0}(\boldsymbol{z}_0, 0, \dots, 0)\| &\le C_m \|\boldsymbol{z}_0\| \sum_{j=1}^{L} \|\boldsymbol{z}_j\|^2 . \end{aligned} \tag{B.19}$$

For the second inequality, the derivative with respect to $(\boldsymbol{z}_1, \dots, \boldsymbol{z}_L)$ vanishes when these coordinates are zero. Its second derivative vanishes when $\boldsymbol{z}_0 = 0$, so Taylor's integral remainder has the displayed factor $\|\boldsymbol{z}_0\|$.

For $\boldsymbol{w} \in \mathbb{R}^2$, set $\widehat{\Phi}_m(\boldsymbol{w}) = \Phi_{m,0}(\boldsymbol{w}, 0, \dots, 0)$. Expanding the numerator and denominator in (B.18) gives

$$\widehat{\Phi}_m(\boldsymbol{w}) = \boldsymbol{w} + \frac{m^2}{24}(w_1^3, w_2^3) - \frac{m}{4} \|\boldsymbol{w}\|^2 \boldsymbol{w} + O(\|\boldsymbol{w}\|^5).$$

Its inverse Taylor polynomial through degree three is therefore

$$\begin{aligned} P_m(\boldsymbol{w}) &= \boldsymbol{w} - \frac{m^2}{24}(w_1^3, w_2^3) + \frac{m}{4} \|\boldsymbol{w}\|^2 \boldsymbol{w}, \\ P_m(\widehat{\Phi}_m(\boldsymbol{w})) &= \boldsymbol{w} + O(\|\boldsymbol{w}\|^5). \end{aligned} \tag{B.20}$$

The symmetric trilinear form defining the cubic term of P_m has operator norm at most $m^2/24 + m/4$.

Use the rotations R_φ, the polynomial Υ, and the real matrix $\boldsymbol{\Omega}$ from (B.13) and (B.14). The transfer matrix satisfies

$$\|\boldsymbol{\Omega}(\omega)\|_{\mathrm{op}} = \frac{2\sin(\omega/2)}{\sqrt{5 - 4\cos\omega}} \le \omega \qquad (0 < \omega \le \pi/2).$$

Fix $c = \alpha = b = 1/64$, and set

$$\theta = \frac{\pi}{2^{L+1}}, \qquad \tau = m^2\theta, \qquad \eta_m = \frac{m\tau}{2}.$$

For sufficiently large m, one has $m \ge 8$ and $\tau \le 1$. Choose $\varepsilon \in (0, 1/m]$ sufficiently small after fixing m. Set $\boldsymbol{\xi} = (\varepsilon/2)(1,0)^\mathsf{T}$, and define

$$\iota_m(\boldsymbol{w}) = \left(\frac{\boldsymbol{\Omega}(\theta)}{\tau} + c\varepsilon^2 I\right)(P_m(\boldsymbol{w}) + \boldsymbol{\xi}) - c\|\boldsymbol{w} + \boldsymbol{\xi}\|^2 (\boldsymbol{w} + \boldsymbol{\xi}). \tag{B.21}$$

The first four players will have identical strategies $p(\boldsymbol{z}^{(t)})$ and expected utilities $1/2 + (\boldsymbol{V}\boldsymbol{\iota}^{(t)})_a$, where $\boldsymbol{\iota}^{(t)} = (\boldsymbol{\iota}_j^{(t)})_{j=0}^{L}$ and

$$\boldsymbol{\iota}_0^{(t)} = \iota_m(\Phi_{m,0}(\boldsymbol{z}^{(t)})),$$

$$\boldsymbol{\iota}_j^{(t)} = -\alpha\Phi_{m,j}(\boldsymbol{z}^{(t)}) + b\Upsilon\left(\Phi_{m,j-1}(\boldsymbol{z}^{(t)}) + \boldsymbol{\xi}\,\mathbb{1}_{\{j=1\}}\right) \qquad (1 \le j \le L). \tag{B.22}$$

To realize these formulas, add $1/2$ to the following pure-payoff deviations. For an own action in block zero, polarize each homogeneous part of ι_m and average its evaluation over subsets of distinct opponents of the corresponding size. Take the own feature's inner product with the resulting vector. For an own action in block $j \ge 1$, take its feature's inner product with the linear term averaged over the three opponents and the bilinear polarization of the quadratic term averaged over their three unordered pairs. Constant terms remain unchanged. Every monomial uses distinct opponents, so independence gives (B.22). An unused action has utility $1/2$. Symmetry among the first four players preserves identical strategies.

The symmetric polarization of $\boldsymbol{w} \mapsto \|\boldsymbol{w}\|^2 \boldsymbol{w}$ sends $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^2$ to

$$\frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle \boldsymbol{w} + \langle \boldsymbol{u}, \boldsymbol{w}\rangle \boldsymbol{v} + \langle \boldsymbol{v}, \boldsymbol{w}\rangle \boldsymbol{u}}{3},$$

so its operator norm is at most one. Every pure feature has norm at most one, and each own action belongs to at most one block. For $i \in [4]$ and pure actions $a_1, \dots, a_6 \in [m]$ with a_i in block zero, (B.20) and $\|\boldsymbol{\Omega}(\theta)\|_{\mathrm{op}} \le \theta$ give

$$\begin{aligned}\left|\mathcal{U}_i(a_1, \dots, a_6) - \frac{1}{2}\right| \le& \left(\frac{1}{m^2} + c\varepsilon^2\right)\left(1 + \frac{m^2}{24} + \frac{m}{4} + \frac{\varepsilon}{2}\right) \\ &+ c\left(1 + \frac{\varepsilon}{2}\right)^3 \\ \le& \frac{65}{64m^2}\frac{m^2}{8} + \frac{1}{32} = \frac{81}{512} < \frac{1}{4}.\end{aligned}$$

Here $m \ge 8$ and $\varepsilon \le 1/m$ give $1 + m^2/24 + m/4 + \varepsilon/2 \le m^2/8$ and $(1 + \varepsilon/2)^3 \le 2$. The other deviations are at most $\alpha + b(1 + \|\boldsymbol{\xi}\|)^2 \le 5/64 < 1/4$, since $\|\boldsymbol{\xi}\| \le 1$. The resulting utilities lie in $[0, 1]$.

At rate η_m, (B.1) and (B.18) give

$$\boldsymbol{z}^{(1)} = 0, \qquad \boldsymbol{\iota}^{(0)} = 0, \qquad \boldsymbol{z}^{(t+1)} = \boldsymbol{z}^{(t)} + \tau\left(2\boldsymbol{\iota}^{(t)} - \boldsymbol{\iota}^{(t-1)}\right). \tag{B.23}$$

We first prove, simultaneously for all $t \ge 1$, that

$$\left\|\boldsymbol{z}_0^{(t)}\right\| \le 3\varepsilon, \qquad \left\|\boldsymbol{z}_j^{(t)}\right\| \le r_j\varepsilon^{2^j} \qquad (1 \le j \le L), \tag{B.24}$$

for constants $r_j > 0$ independent of ε. Define

$$q = \frac{1 - 2\tau\alpha + \sqrt{1 + 4(\tau\alpha)^2}}{2} \in (0, 1).$$

The linear companion map $(\boldsymbol{z}, \boldsymbol{y}) \mapsto ((1 - 2\tau\alpha)\boldsymbol{z} + \tau\alpha\boldsymbol{y}, \boldsymbol{z})$ has norm q in $\max\{\|\boldsymbol{z}\|, q\|\boldsymbol{y}\|\}$. Moreover, $q \ge 1/\sqrt{2}$ and $1 - q \ge \tau\alpha/2$.

Temporarily suppose $\left\|\boldsymbol{z}_0^{(t)}\right\| \le 4\varepsilon$ and the other bounds in (B.24) hold. Choose ε so that $\sum_{j=1}^{L} r_j^2\varepsilon^{2^{j+1}} \le \varepsilon^2$; then $\left\|\boldsymbol{z}^{(t)}\right\|^2 \le 17\varepsilon^2$. Set $r_0 = 9/2$. For $1 \le j \le L$ and $t \ge 2$, define $v_j^{(t)} = \max\left\{\left\|\boldsymbol{z}_j^{(t)}\right\|, q\left\|\boldsymbol{z}_j^{(t-1)}\right\|\right\}$. By (B.19) and (B.22),

$$v_j^{(t+1)} \le (q + 68\tau\alpha C_m\varepsilon^2)v_j^{(t)} + 12\tau b r_{j-1}^2\varepsilon^{2^j}. \tag{B.25}$$

Here the preceding moment, including the shift when $j = 1$, has norm at most $2r_{j-1}\varepsilon^{2^{j-1}}$. For $\varepsilon^2 \le 1/(272C_m)$, the coefficient of $v_j^{(t)}$ is at most $(1 + q)/2$. Choose successively $r_j = 1 + 48\tau b r_{j-1}^2/(1 - q)$. The recurrence preserves each bound with strict slack. At the first update, $\boldsymbol{z}_1^{(2)} = (\tau b\varepsilon^2/2)(1, 0)^{\mathsf{T}}$ and $\boldsymbol{z}_j^{(2)} = 0$ for $j \ge 2$, which satisfy the same bounds.

To close the block-zero bound, note that the provisional estimates give $\sum_{j=1}^{L} \|\boldsymbol{z}_j\|^2 \le O(\varepsilon^4)$. Consequently, $\Phi_{m,0}(\boldsymbol{z}) = \widehat{\Phi}_m(\boldsymbol{z}_0) + O(\varepsilon^5)$. For bounded $\widehat{\boldsymbol{z}}$ and $\boldsymbol{z}_0 = \varepsilon\widehat{\boldsymbol{z}} - \boldsymbol{\xi}$, (B.20) and (B.21) give

$$\iota_m(\Phi_{m,0}(\boldsymbol{z})) = \frac{\varepsilon}{\tau}\boldsymbol{\Omega}(\theta)\widehat{\boldsymbol{z}} + c\varepsilon^3(1 - \|\widehat{\boldsymbol{z}}\|^2)\widehat{\boldsymbol{z}} + O(\varepsilon^4).$$

The error is uniform over all other blocks satisfying the provisional bounds.

Set $\widehat{\boldsymbol{z}}^{(t)} = (\boldsymbol{z}_0^{(t)} + \boldsymbol{\xi})/\varepsilon$, $\boldsymbol{R}_{\mathrm{f}} = (R_\theta - I)(2R_\theta - I)^{-1}$, and $\mu = \tau c\varepsilon^2$. These real matrices commute and satisfy

$$\begin{aligned}\boldsymbol{\Omega}(\theta) =& R_\theta \boldsymbol{R}_{\mathrm{f}}, \qquad I + 2\boldsymbol{\Omega}(\theta) = R_\theta + \boldsymbol{R}_{\mathrm{f}},\\ \kappa =& \frac{1}{2}\operatorname{tr}\left[R_{-\theta}(R_\theta - \boldsymbol{R}_{\mathrm{f}})^{-1}(2R_\theta - I)\right].\end{aligned}$$

For sufficiently large m,

$$\|\boldsymbol{R}_{\mathrm{f}}\|_{\mathrm{op}} \le \frac{1}{4}, \qquad \frac{2}{5} \le \left\|\frac{1}{2}(R_\theta - \boldsymbol{R}_{\mathrm{f}})^{-1}R_\theta(1,0)^{\mathsf{T}}\right\| \le \frac{3}{5}, \qquad \kappa \ge \frac{1}{2}.$$

The matrix inside the trace has symmetric part κI. The normalized recurrence is

$$\begin{aligned}\widehat{\boldsymbol{z}}^{(t+1)} =& (R_\theta + \boldsymbol{R}_{\mathrm{f}})\widehat{\boldsymbol{z}}^{(t)} - R_\theta \boldsymbol{R}_{\mathrm{f}}\widehat{\boldsymbol{z}}^{(t-1)}\\ &+ \mu\left[2(1 - \left\|\widehat{\boldsymbol{z}}^{(t)}\right\|^2)\widehat{\boldsymbol{z}}^{(t)} - (1 - \left\|\widehat{\boldsymbol{z}}^{(t-1)}\right\|^2)\widehat{\boldsymbol{z}}^{(t-1)}\right] + O(\mu\varepsilon),\end{aligned}$$

with $\widehat{\boldsymbol{z}}^{(1)} = (1,0)^{\mathsf{T}}/2$ and $\widehat{\boldsymbol{z}}^{(2)} = (I + 2\boldsymbol{\Omega}(\theta))(1,0)^{\mathsf{T}}/2 + O(\mu)$. For $t \ge 2$, define

$$\begin{aligned}\boldsymbol{w}^{(t)} =& (R_\theta - \boldsymbol{R}_{\mathrm{f}})^{-1}(\widehat{\boldsymbol{z}}^{(t)} - \boldsymbol{R}_{\mathrm{f}}\widehat{\boldsymbol{z}}^{(t-1)}),\\ \boldsymbol{v}^{(t)} =& (R_\theta - \boldsymbol{R}_{\mathrm{f}})^{-1}(R_\theta\widehat{\boldsymbol{z}}^{(t-1)} - \widehat{\boldsymbol{z}}^{(t)}).\end{aligned}$$

Since $\widehat{\boldsymbol{z}}^{(t)} = R_\theta\boldsymbol{w}^{(t)} + \boldsymbol{R}_{\mathrm{f}}\boldsymbol{v}^{(t)}$ and $\widehat{\boldsymbol{z}}^{(t-1)} = \boldsymbol{w}^{(t)} + \boldsymbol{v}^{(t)}$, substitution gives on bounded sets

$$\begin{aligned}\boldsymbol{w}^{(t+1)} =& R_\theta\boldsymbol{w}^{(t)} + \mu(R_\theta - \boldsymbol{R}_{\mathrm{f}})^{-1}(2R_\theta - I)(1 - \left\|\boldsymbol{w}^{(t)}\right\|^2)\boldsymbol{w}^{(t)} + O(\mu\left\|\boldsymbol{v}^{(t)}\right\| + \mu\varepsilon),\\ \boldsymbol{v}^{(t+1)} =& \boldsymbol{R}_{\mathrm{f}}\boldsymbol{v}^{(t)} + O(\mu).\end{aligned}$$

The initial values are

$$\begin{aligned}\boldsymbol{w}^{(2)} =& \tfrac{1}{2}(R_\theta - \boldsymbol{R}_{\mathrm{f}})^{-1}R_\theta(1,0)^{\mathsf{T}} + O(\mu),\\ \boldsymbol{v}^{(2)} =& -\tfrac{1}{2}(R_\theta - \boldsymbol{R}_{\mathrm{f}})^{-1}\boldsymbol{R}_{\mathrm{f}}(1,0)^{\mathsf{T}} + O(\mu).\end{aligned}$$

After $n_0 = \lceil \log(1/\mu)/\log 4 \rceil$ steps, $\left\|\boldsymbol{v}^{(t)}\right\| \le O(\mu)$, while $\left\|\boldsymbol{w}^{(t)}\right\|$ has changed by at most $O(\mu n_0) = o(1)$. A first-exit argument on $\left\|\boldsymbol{w}^{(t)}\right\| \le 3/2$ and $\left\|\boldsymbol{v}^{(t)}\right\| \le 1$ therefore gives $1/3 \le \left\|\boldsymbol{w}^{(n_0+2)}\right\| \le 3/4$. Thereafter, for $1/4 \le \left\|\boldsymbol{w}^{(t)}\right\| \le 3/2$,

$$\left\|\boldsymbol{w}^{(t+1)}\right\| = \left\|\boldsymbol{w}^{(t)}\right\| + \mu\kappa\left\|\boldsymbol{w}^{(t)}\right\|(1 - \left\|\boldsymbol{w}^{(t)}\right\|^2) + O(\mu\varepsilon). \tag{B.26}$$

For sufficiently small ε, the polynomial $r \mapsto r + \mu\kappa r(1 - r^2)$ is increasing on $[1/4, 3/2]$. Its inward drift at the endpoints exceeds the uniform error in (B.26), so this interval is invariant. For a fixed constant $C_{m,0} > 0$ bounding that error,

$$\left|\left\|\boldsymbol{w}^{(t+1)}\right\| - 1\right| \le \left(1 - \frac{5\kappa\mu}{16}\right)\left|\left\|\boldsymbol{w}^{(t)}\right\| - 1\right| + C_{m,0}\mu\varepsilon.$$

Indeed, $r + \mu\kappa r(1 - r^2) - 1 = (r - 1)(1 - \mu\kappa r(r + 1))$ and $r(r + 1) \ge 5/16$. It follows that $\left\|\boldsymbol{w}^{(t)}\right\| = 1 + O(\varepsilon)$ after a finite prefix. Throughout this argument, including that prefix, $\left\|\boldsymbol{z}_0^{(t)}\right\| \le 3\varepsilon$, strictly below the provisional bound. Together with (B.25), this rules out the first exit of any block and proves (B.24).

Lifting the polar angle of $R_\theta \boldsymbol{w}^{(t)}$ along the tail gives real phases $\vartheta^{(t)}$ such that

$$\begin{aligned}\Phi_{m,0}(\boldsymbol{z}^{(t)}) + \boldsymbol{\xi} =& \varepsilon R_{\vartheta^{(t)}}(1,0)^{\mathsf{T}} + O(\varepsilon^2),\\ \vartheta^{(t+1)} - \vartheta^{(t)} =& \theta + O(\varepsilon^2).\end{aligned} \tag{B.27}$$

Both errors are uniform in time.

Set $\boldsymbol{\nu}_0 = (1,0)^{\mathsf{T}}$ and, for $1 \le j \le L$, define

$$\boldsymbol{\nu}_j = \left(\alpha I + \frac{\boldsymbol{\Omega}(2^j\theta)}{\tau}\right)^{-1} b\Upsilon(\boldsymbol{\nu}_{j-1}).$$

The coefficient of J in $\boldsymbol{\Omega}(\omega)$ is positive for $0 < \omega \le \pi/2$, so the inverse exists. By (B.13), every $\boldsymbol{\nu}_j$ is nonzero. We claim that, for $1 \le j \le L$,

$$\Phi_{m,j}(\boldsymbol{z}^{(t)}) = \varepsilon^{2^j} R_{2^j\vartheta^{(t)}}\boldsymbol{\nu}_j + O(\varepsilon^{2^j+1}) + o_t(1). \tag{B.28}$$

Here $o_t(1)$ tends to zero for fixed ε. By (B.19) and (B.24), $\Phi_{m,j}(\boldsymbol{z}^{(t)}) - \boldsymbol{z}_j^{(t)} = O(\varepsilon^{2^j+2})$. If the preceding block has the claimed expansion, its quadratic forcing has leading coefficient $b\Upsilon(\boldsymbol{\nu}_{j-1})$ and error $O(\varepsilon^{2^j+1}) + o_t(1)$. For $j = 1$, (B.27) gives this conclusion with the required shift $\boldsymbol{\xi}$. Inserting $\varepsilon^{2^j} R_{2^j\vartheta^{(t)}}\boldsymbol{\nu}_j$ into (B.23) cancels the leading terms because

$$\begin{aligned}-\alpha\boldsymbol{\nu}_j + b\Upsilon(\boldsymbol{\nu}_{j-1}) =& \frac{\boldsymbol{\Omega}(2^j\theta)}{\tau}\boldsymbol{\nu}_j,\\ R_\omega - I =& \boldsymbol{\Omega}(\omega)(2I - R_{-\omega}).\end{aligned}$$

The phase error in (B.27) contributes $O(\varepsilon^{2^j+2})$. The difference from the proposed sequence satisfies the linear companion recurrence with norm $q < 1$ and forcing $O(\varepsilon^{2^j+1}) + o_t(1)$. Summing this geometric recurrence proves (B.28) by induction.

Players five and six also have all m actions. If $i \in \{5,6\}$ and i' is the other such player, set

$$\mathcal{U}_i(a_i, a_{i'}, a_1) = \frac{1}{2} + \boldsymbol{\zeta}_{0,a_i} \cdot \left[-\alpha\boldsymbol{\zeta}_{0,a_{i'}} + b\boldsymbol{\zeta}_{L,a_1}\right].$$

Their pure utilities lie in $[0,1]$. They remain identical, with strategies $p(\boldsymbol{z}_*^{(t)}, 0, \dots, 0)$, where $\boldsymbol{z}_*^{(t)} \in \mathbb{R}^2$, moments $\widehat{\Phi}_m(\boldsymbol{z}_*^{(t)})$, and feedback $\boldsymbol{\iota}_*^{(t)} = -\alpha\widehat{\Phi}_m(\boldsymbol{z}_*^{(t)}) + b\Phi_{m,L}(\boldsymbol{z}^{(t)})$. In particular, the denominator in their softmax includes all m actions. The companion estimate gives $\sup_t \left\|\boldsymbol{z}_*^{(t)}\right\| \le O(\varepsilon^{2^L})$. Define

$$\boldsymbol{\nu}_* = \left(\alpha I + \frac{\boldsymbol{\Omega}(\pi/2)}{\tau}\right)^{-1} b\boldsymbol{\nu}_L \neq 0.$$

Applying the preceding recurrence calculation without squaring the forcing gives, since $2^L\theta = \pi/2$,

$$\begin{aligned}\widehat{\Phi}_m(\boldsymbol{z}_*^{(t)}) =& \varepsilon^{2^L} R_{2^L\vartheta^{(t)}}\boldsymbol{\nu}_* + O(\varepsilon^{2^L+1}) + o_t(1),\\ \boldsymbol{\iota}_*^{(t)} =& \frac{\boldsymbol{\Omega}(\pi/2)}{\tau}\varepsilon^{2^L} R_{2^L\vartheta^{(t)}}\boldsymbol{\nu}_* + O(\varepsilon^{2^L+1}) + o_t(1).\end{aligned}$$

The uniform comparator earns $1/2$ at every timestep. The identities $\boldsymbol{\Omega}(\pi/2) = (-I + 3J)/5$ and $\boldsymbol{z} \cdot J\boldsymbol{z} = 0$ yield

$$-\widehat{\Phi}_m(\boldsymbol{z}_*^{(t)}) \cdot \boldsymbol{\iota}_*^{(t)} = \frac{\|\boldsymbol{\nu}_*\|^2}{5\tau}\varepsilon^{2^{L+1}} + O(\varepsilon^{2^{L+1}+1}) + o_t(1).$$

Choose $\varepsilon > 0$ sufficiently small after all preceding constants. The comparator advantage has strictly positive lower limit, so either of these players has unbounded cumulative regret.

For every $\eta \ge \eta_m$, replace each utility $\mathcal{U}_i$ by $1/2 + (\eta_m/\eta)(\mathcal{U}_i - 1/2)$. The utilities remain in $[0,1]$, the strategy trajectory is unchanged, and regret is multiplied by $\eta_m/\eta > 0$. Its supremum over time remains infinite. For $0 < \eta < \eta_m$, give one player one action with constant utility

one and $m-1$ actions with constant utility zero, and give the other five players constant utilities. Set $r(s) = (m-1)/(m-1+e^s)$ for $s \geq 0$. This function is nonnegative and decreasing, and $\int_0^\infty r(s)\,\mathrm{d}s = \log m$. Thus

$$\mathrm{Reg}_1(T) = r(0) + \sum_{t=2}^{T} r(\eta t) \geq \frac{1}{\eta}\int_0^{\eta(T+1)} r(s)\,\mathrm{d}s - r(\eta).$$

Taking the supremum over T gives

$$\sup_T \mathrm{Reg}_1(T) \geq \frac{\log m}{\eta} - 1 > \frac{\log m}{\eta_m} - 1 = \frac{2}{\pi}\frac{2^{\lfloor m/4 \rfloor}\log m}{m^3} - 1.$$

□

C THE CLOSED-FORM REGULARIZED RESPONSE

Lemma 4.1. *For every $i \in [N]$, $1 \leq M_i \leq e$ and*

$$\mathrm{dom}\,\psi_i = \left\{ \boldsymbol{y} \in \mathbb{R}^{m_i}_{\geq 0} : \|\boldsymbol{y}\|_{p_i/(p_i-1)} \leq 1/M_i \right\}.$$

The range of ψ_i is r_i. For every $\boldsymbol{\theta} \in \mathbb{R}^{m_i}$, the maximizer $\boldsymbol{y}_i(\boldsymbol{\theta})$ in (4.1) is unique and belongs to the interior of $\mathrm{dom}\,\psi_i$. Moreover, $0 < \lambda_i(\boldsymbol{\theta}) < 1$ and $\nabla\psi_i(\boldsymbol{y}_i(\boldsymbol{\theta})) = \boldsymbol{\theta}$. In addition, $\boldsymbol{y}_i(\mathbf{0}) = \mathbf{1}/(2m_i)$. Defining $\mathrm{asinh}(z) \coloneqq \log\left(z + \sqrt{1+z^2}\right)$, we have, for every $a_i \in [m_i]$,

$$x_i(\boldsymbol{\theta})(a_i) = \frac{\exp\left(p_i\,\mathrm{asinh}(\theta(a_i)/(2r_i)) - \frac{1}{2}\log(1+(\theta(a_i)/(2r_i))^2)\right)}{\sum_{b_i=1}^{m_i}\exp\left(p_i\,\mathrm{asinh}(\theta(b_i)/(2r_i)) - \frac{1}{2}\log(1+(\theta(b_i)/(2r_i))^2)\right)}. \tag{4.2}$$

Proof. Fix a player $i \in [N]$ and suppress its index. Write $m = m_i$, $p = p_i$, $M = M_i$, $r = r_i$, and $\psi = \psi_i$, and let ψ^* denote the Fenchel conjugate of ψ. Define $h(s) = (s + \sqrt{1+s^2})/2$ for $s \in \mathbb{R}$. All coordinate sums in this proof run over $[m]$. Since $p \geq \log m$ and $m \geq 2$, one has $1 \leq M = m^{1/p} \leq e$. The joint feasible set in the definition of ψ is compact and convex, and the objective is continuous and convex. Thus partial minimization gives a proper closed convex function. For fixed $\boldsymbol{w} = \boldsymbol{y} + \boldsymbol{z}$, coordinatewise maximization uses

$$\max_{0\leq t\leq 1}\left\{ st + \sqrt{t(1-t)} \right\} = h(s), \qquad t = h'(s) = \frac{1}{2}\left(1 + \frac{s}{\sqrt{1+s^2}}\right).$$

The maximizer is unique, as follows by differentiation and strict concavity. Hölder's inequality and its equality condition therefore give, with $\boldsymbol{h} \in \mathbb{R}^m$ defined by $h_a = h(\theta(a)/(2r))$,

$$\psi^*(\boldsymbol{\theta}) = \max_{\substack{\boldsymbol{w}\geq 0 \\ \|\boldsymbol{w}\|_{p/(p-1)}\leq 1/M}} 2r\,\langle \boldsymbol{w}, \boldsymbol{h}\rangle = \frac{2r}{M}\,\|\boldsymbol{h}\|_p\,,$$

$$w(a) = \frac{1}{M}\left(\frac{h_a}{\|\boldsymbol{h}\|_p}\right)^{p-1}, \qquad y(a) = w(a)h'(\theta(a)/(2r)).$$

The maximizing $\boldsymbol{y}$ is unique and equals $\nabla\psi^*(\boldsymbol{\theta})$, so it is the response in (4.1). Since $0 < h' < 1$, all response coordinates are positive and $\|M\boldsymbol{y}\|_{p/(p-1)} < 1$. Moreover, $\sum_a w(a) \leq 1$ by Hölder's inequality, so $0 < \lambda = \sum_a y(a) < 1$. The identities $h(s) = e^{\mathrm{asinh}\,s}/2$ and $h'(s) = h(s)/\sqrt{1+s^2}$ give the asserted normalized response.

The domain consists exactly of the nonnegative vectors of $\ell_{p/(p-1)}$ norm at most $1/M$. On this domain,

$$-r \leq -r\sum_a (y(a) + z(a)) \leq -2r\sum_a \sqrt{y(a)z(a)} \leq 0.$$

At zero score the optimizer has $y(a) = z(a) = 1/(2m)$; thus $\lambda(\mathbf{0}) = 1/2$, $x(\mathbf{0})(a) = 1/m$, and $\psi(\boldsymbol{y}(\mathbf{0})) = -r$. Also $\psi(\mathbf{0}) = 0$, so the range is exactly r.

Differentiating the positive ℓ_p norm shows that, for every nonzero $\boldsymbol{v} \in \mathbb{R}^m$,

$$\boldsymbol{v}^\top \nabla^2 \psi^*(\boldsymbol{\theta}) \boldsymbol{v} \geq \frac{1}{2rM} \sum_a \left(\frac{h_a}{\|\boldsymbol{h}\|_p} \right)^{p-1} h''(\theta(a)/(2r)) v(a)^2 > 0.$$

To verify smoothness of ψ in the interior of its domain, fix an interior feasible $\boldsymbol{y}$ and a ball of radius $a > 0$ about $\boldsymbol{y}$ contained in the domain. The preceding range bound implies

$$\psi^*(\boldsymbol{\theta}) - \langle \boldsymbol{\theta}, \boldsymbol{y} \rangle \geq a \|\boldsymbol{\theta}\|_2 .$$

Consequently this function has a unique minimizer. Its first-order condition is $\nabla \psi^*(\boldsymbol{\theta}) = \boldsymbol{y}$, and the inverse function theorem gives the smooth inverse $\nabla \psi(\boldsymbol{y}) = \boldsymbol{\theta}$. □

D Proofs for general filters

D.1 Formal statement and proof

For a filter $F(z) = \sum_{s \geq 0} f_s z^s$ satisfying Assumption 5.1, let k be as in Section 5 and define $\kappa_0 \coloneqq \sum_{s \geq 0} |f_s| \geq 1$. The series converges at the real point R from that assumption, so it defines an analytic function on $\{z \in \mathbb{C} : |z| < R\}$. All filter maxima below are over complex z. For $k \geq 1$, use τ, ρ_τ from Section 5.3 and F_h from (5.2), and define

$$\begin{aligned} R_\tau &\coloneqq \sqrt{1 + \frac{1}{2\tau}} < R, \\ K_h^F(z) &\coloneqq \frac{F_h(z)}{1 - z} \qquad (1 \leq h \leq k), \\ K_\star^F(z) &\coloneqq \frac{1 - (1 - z)F(z)}{1 - \rho_\tau z}. \end{aligned}$$

These functions are analytic on a neighborhood of $\{z : |z| \leq R_\tau\}$; the singularities at $z = 1$ are removable. Set $\mathsf{K}_h^F \coloneqq K_h^F(\mathsf{L})$, $\mathsf{K}_\star^F \coloneqq K_\star^F(\mathsf{L})$, and $\mathsf{C} \coloneqq I - \mathsf{DF}$. The other two filter norms are

$$\begin{aligned} \kappa_1 &\coloneqq \max_{1 \leq h \leq k} \max_{|z| \leq R_\tau} \left| K_h^F(z) \right|, \\ \kappa_2 &\coloneqq \max \left\{ 1, \frac{1}{\tau} \max_{|z| \leq R_\tau} \left| K_\star^F(z) \right| \right\}. \end{aligned}$$

The following theorem retains the explicit constants used in Theorem 4.2 and corollary D.6.

Theorem D.1. *Suppose Assumption 5.1 holds, and use k, κ_0 defined above. If $k = 0$, the update in* (4.7) *and* (5.1)*, with $0 < \eta \leq 1/(20\kappa_0)$, satisfies, for every $i \in [N]$ and $T \geq 1$,*

$$[\mathrm{Reg}_i(T)]_+ \leq M_i \left(\frac{r_i}{\eta} + 3\eta \kappa_0^2 T \right).$$

For $k \geq 1$, fix τ, κ_1, κ_2 as above. For a common learning rate $\eta > 0$, define the following quantities subject to the indicated inequalities:

$$\begin{aligned} q &\coloneqq 5\eta^2 \tau^2 \kappa_2^2 < 1, \qquad \alpha_0 \coloneqq \frac{6\kappa_1^2}{1 - q}, \\ \alpha_1 &\coloneqq \frac{\kappa_1^2}{1 - q} \left[10N + (4000 + 10000N) (2\kappa_0 - 1)^2 \eta^2 \tau^2 \right], \qquad \beta \coloneqq 3\eta^2 N \alpha_1 < \frac{1}{2}. \end{aligned} \tag{D.1}$$

For every integer $T \geq 1$, set

$$\alpha_{0,T} \coloneqq \alpha_0 + 10\kappa_1^2 (2\kappa_0 - 1)^2 \eta^2 (N - k) q^{k-1} T.$$

If all players use (4.7) *and* (5.1) *with the common filter F and learning rate η satisfying $\eta(2\kappa_0 - 1) \leq 1/(20\tau)$, then every player $i \in [N]$ satisfies, simultaneously for every $T \geq 1$,*

$$[\mathrm{Reg}_i(T)]_+ \leq \frac{M_i}{\eta} \left(r_i + \frac{\beta}{1/2 - \beta} \frac{1}{N} \sum_j r_j \right) + \frac{3M_i \eta \alpha_{0,T}}{1 - 2\beta}. \tag{D.2}$$

In particular, every $0 < \eta \le (1000\tau(2\kappa_0 - 1)\kappa_1\kappa_2)^{-1}$ *satisfies these learning-rate conditions and gives*

$$[\mathrm{Reg}_i(T)]_+ \le \frac{2M_i}{\eta}\left(r_i + \frac{1}{N}\sum_{j=1}^{N} r_j\right) + 60M_i\kappa_1^2(2\kappa_0 - 1)^2(N-k)\eta^3 q^{k-1}T. \tag{D.3}$$

Proof. For $k = 0$, the conclusion is (D.24). Assume $k \ge 1$ for the remainder of the proof.

Use the operators $\mathsf{F}_h, \mathsf{K}_h^F, \mathsf{K}_\star^F$ defined above, and recall that $\mathsf{C} = I - \mathsf{DF}$.

Fix $T \ge 1$. Unless another range is displayed, sums over time run from 1 to T and sums over players run over $[N]$. Use the coordinate norms and movement sums $\mathcal{P}_{i,T}, \mathcal{P}_T$ from Section 5.3; operator norms have the indicated input and output norms. Extend $\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{\lambda}_i, \boldsymbol{\theta}_i$ constantly to nonpositive times using their values at zero score. Let $\boldsymbol{J}_{\boldsymbol{x}_i}(\boldsymbol{\theta})$ be the Jacobian of the normalized response map $\boldsymbol{x}_i$ at $\boldsymbol{\theta}$. For every $i \in [N]$ and integer $t \ge 1$, define

$$\begin{aligned} P_i^{(t)} &:= D_{\psi_i}(\boldsymbol{y}_i^{(t)}, \boldsymbol{y}_i^{(t-1)}), \\ \boldsymbol{d}_i &:= \mathsf{C}\boldsymbol{g}_i, \qquad \overline{\boldsymbol{J}}_i^{(t)} := \eta\int_0^1 \boldsymbol{J}_{\boldsymbol{x}_i}\left(\boldsymbol{\theta}_i^{(t-1)} + a(\boldsymbol{\theta}_i^{(t)} - \boldsymbol{\theta}_i^{(t-1)})\right)\mathrm{d}a. \end{aligned} \tag{D.4}$$

At nonpositive times use $\overline{\boldsymbol{J}}_i^{(t)} = \eta\boldsymbol{J}_{\boldsymbol{x}_i}(0)$ and $P_i^{(t)} = 0$. The score identity is $\mathsf{D}\boldsymbol{\theta}_i = \eta\mathsf{C}\boldsymbol{g}_i = \eta\boldsymbol{d}_i$, including time 1, because $f_0 = 1$ and $\boldsymbol{\theta}_i^{(1)} = 0$. Hence $\mathsf{D}\boldsymbol{x}_i = \overline{\boldsymbol{J}}_i\boldsymbol{d}_i$.

For a scalar filter $K(z) = \sum_{s\ge0} a_s z^s$, write $\|K\|_{\ell^1} := \sum_{s\ge0}|a_s|$. Since $f_0 = 1$, we have $\kappa_0 \ge 1$ and $\|I - \mathsf{F}\|_{\ell^1} = \kappa_0 - 1$. The identity $\mathsf{C} = \mathsf{L} + \mathsf{D}(I - \mathsf{F})$ gives

$$\|\mathsf{C}\|_{\ell^1} \le 1 + 2\|I - \mathsf{F}\|_{\ell^1} = 2\kappa_0 - 1.$$

Write $\epsilon = \eta(2\kappa_0 - 1)$. Since $K_1^F(0) = 1$, $\kappa_1 \ge 1$, and $\epsilon \le 1/(20\tau) \le 1/40$. Every score increment has ℓ_∞ norm at most ϵ, and $\left\|\boldsymbol{e}_i^{(t)}\right\|_\infty \le \kappa_0$ and $\left\|\boldsymbol{d}_i^{(t)}\right\|_\infty \le 2\kappa_0 - 1$. The following estimates apply to these score increments.

Lemma D.2. *Let* $0 < \epsilon \le 1/25$. *For every* $i \in [N]$, *consider scores* $\boldsymbol{\theta}_i^{(t)}$ *satisfying* $\left\|\boldsymbol{\theta}_i^{(t)} - \boldsymbol{\theta}_i^{(t-1)}\right\|_\infty \le \epsilon$ *for* $t \in [T]$, *with constant extension to* $t \le 0$. *Use the responses and quantities in* (D.4). *For every* $t \in [T]$, $a_i \in [m_i]$, *and* $\boldsymbol{v} \in \mathbb{R}^{m_i}$,

$$\begin{aligned} \left|\log\frac{\lambda_i^{(t)}}{\lambda_i^{(t-1)}}\right| &\le \epsilon, \qquad \left|\log\frac{x_i^{(t)}(a_i)}{x_i^{(t-1)}(a_i)}\right| \le 2\epsilon, \\ \sqrt{\lambda_i^{(t)}}\left\|(\mathsf{D}\boldsymbol{x}_i)^{(t)}\right\|^2_{(\boldsymbol{x}_i^{(t)})^{-1}} &\le 5P_i^{(t)}, \\ \left\|\overline{\boldsymbol{J}}_i^{(t)}\boldsymbol{v}\right\|^2_{(\boldsymbol{x}_i^{(t)})^{-1}} &\le 5\eta^2\lambda_i^{(t)}\|\boldsymbol{v}\|^2_{\boldsymbol{x}_i^{(t)}}, \\ \sum_{t=1}^{T}\left\|\overline{\boldsymbol{J}}_i^{(t)} - \overline{\boldsymbol{J}}_i^{(t-1)}\right\|^2_{\infty\to(\boldsymbol{x}_i^{(t)})^{-1}} &\le 60\eta^2\mathcal{P}_{i,T}. \end{aligned} \tag{D.5}$$

For every $t \in [T]$, *set* $b_{i,t} := \left\|\sqrt{\boldsymbol{x}_i^{(t)}} - \sqrt{\boldsymbol{x}_i^{(t-1)}}\right\|_2$, *where square roots are coordinatewise. Then*

$$b_{i,t} \le \epsilon e^{\epsilon/2}\sqrt{\lambda_i^{(t)}}, \qquad \sqrt{\lambda_i^{(t)}}b_{i,t}^2 \le e^{11\epsilon/4}P_i^{(t)}. \tag{D.6}$$

Its proof is in Section E.

Define the lifted regret $\mathcal{R}_{i,T} = \max_{\widehat{\boldsymbol{y}}_i \in \mathrm{dom}\,\psi_i}\sum_{t=1}^{T}\left\langle \boldsymbol{g}_i^{(t)}, \widehat{\boldsymbol{y}}_i - \boldsymbol{y}_i^{(t)}\right\rangle$. Centering gives $\left\langle \boldsymbol{g}_i^{(t)}, \boldsymbol{y}_i^{(t)}\right\rangle = 0$. The feasible comparators $\mathbf{0}$ and every pure strategy scaled by $1/M_i$ therefore give $\mathcal{R}_{i,T} \ge 0$ and $[\mathrm{Reg}_i(T)]_+ \le M_i\mathcal{R}_{i,T}$. The next lemma bounds this lifted regret.

Lemma D.3. *Fix $i \in [N]$. For the update in* (5.1)*, suppose $2\eta \left\| \boldsymbol{e}_i^{(t)} \right\|_\infty \le 1/10$ for every $t \in [T]$. Then*

$$\mathcal{R}_{i,T} \le \frac{r_i}{\eta} + 3\eta \sum_{t=1}^{T} (\lambda_i^{(t)})^{3/2} \left\| (\mathsf{F}\boldsymbol{g}_i)^{(t)} \right\|^2_{\boldsymbol{x}_i^{(t)}} - \frac{\mathcal{P}_{i,T}}{2\eta}. \tag{D.7}$$

The proof is in Section E. Since $2\eta \left\| \boldsymbol{e}_i^{(t)} \right\|_\infty \le 2\epsilon \le 1/10$, Lemma D.3 applies.

Weighted filters and commutators. On a finite-dimensional coordinate space, let $\omega_{t,a} > 0$ for every coordinate a and $t \in [T]$, and define

$$\|\boldsymbol{v}\|^2_{H_t} \coloneqq \sum_a \omega_{t,a} \left| v(a) \right|^2,$$

$$\|\boldsymbol{v}\|^2_{\ell^2(H)} \coloneqq \sum_{t=1}^{T} \left\| \boldsymbol{v}^{(t)} \right\|^2_{H_t} \qquad \text{for } \boldsymbol{v} = (\boldsymbol{v}^{(t)})_{t=1}^T.$$

Suppose the weights satisfy

$$\left| \log \frac{\omega_{t,a}}{\omega_{t-1,a}} \right| \le \frac{1}{5\tau} \qquad \text{for every coordinate } a \text{ and } 2 \le t \le T.$$

For complex vectors, use the same weighted norms on the complexified spaces. Define the finite backward shift $\mathsf{L}_{\mathrm{out}}$ on $\ell^2(H)$ by $(\mathsf{L}_{\mathrm{out}}\boldsymbol{v})^{(1)} = \boldsymbol{0}$ and $(\mathsf{L}_{\mathrm{out}}\boldsymbol{v})^{(t)} = \boldsymbol{v}^{(t-1)}$ for $2 \le t \le T$. Then

$$\|\mathsf{L}_{\mathrm{out}}\|^2 \le e^{1/(5\tau)} \le 1 + \frac{2}{9\tau}, \qquad R_\tau^2 - \|\mathsf{L}_{\mathrm{out}}\|^2 \ge \frac{5}{18\tau} > \frac{1}{4\tau}. \tag{D.8}$$

For every scalar function K analytic on a neighborhood of $\{z : |z| \le R_\tau\}$, von Neumann's inequality applied to $\mathsf{L}_{\mathrm{out}}/R_\tau$ gives $\|K(\mathsf{L}_{\mathrm{out}})\| \le \max_{|z|=R_\tau} |K(z)|$. Consequently K_h^F and $\mathsf{K}_\star^F$ have operator norms at most κ_1 and $\tau\kappa_2$, respectively, in every such space.

We use the following commutator bound.

Lemma D.4. *Let H_t be coordinate Hilbert norms for $t \in [T]$, with the finite backward shift $\mathsf{L}_{\mathrm{out}}$ on $\ell^2(H)$. Let $R > \max\{1, \|\mathsf{L}_{\mathrm{out}}\|\}$ and let K have real Taylor coefficients and be analytic near $\{z : |z| \le R\}$. For real linear maps $\boldsymbol{J}^{(t)}$ from a fixed coordinate space into H_t, $t \in [T]$, set $\boldsymbol{J}^{(0)} = \boldsymbol{J}^{(1)}$. Write $\boldsymbol{J}$ for their pointwise action, set $\Delta\boldsymbol{J}^{(t)} \coloneqq \boldsymbol{J}^{(t)} - \boldsymbol{J}^{(t-1)}$, and define $[K, \boldsymbol{J}]\boldsymbol{v} \coloneqq K(\boldsymbol{J}\boldsymbol{v}) - \boldsymbol{J}(K\boldsymbol{v})$. For every real input sequence $\boldsymbol{v}$ extended by zero outside $[T]$, with $\|\boldsymbol{v}\|_\infty \coloneqq \max_{t\in[T]} \left\| \boldsymbol{v}^{(t)} \right\|_\infty$,*

$$\|[K, \boldsymbol{J}]\boldsymbol{v}\|^2_{\ell^2(H)} \le \frac{\pi R^2 \max_{|z|=R} |K(z)|^2}{2(R^2 - \|\mathsf{L}_{\mathrm{out}}\|^2)(R^2 - 1)} \|\boldsymbol{v}\|^2_\infty \sum_{t=1}^{T} \left\| \Delta\boldsymbol{J}^{(t)} \right\|^2_{\infty \to H_t}. \tag{D.9}$$

Let L_{in} be the finite backward shift on input sequences, with the same zero boundary convention. The proof is in Section E and uses the following resolvent identity with the radius R in Lemma D.4:

$$[K, \boldsymbol{J}] = \frac{1}{2\pi \mathrm{i}} \oint_{|z|=R} K(z)(zI - \mathsf{L}_{\mathrm{out}})^{-1}(\mathsf{L}_{\mathrm{out}}\boldsymbol{J} - \boldsymbol{J}\mathsf{L}_{\mathrm{in}})(zI - \mathsf{L}_{\mathrm{in}})^{-1}\, dz, \tag{D.10}$$

Its middle operator maps $\boldsymbol{v}$ to $-\Delta\boldsymbol{J}^{(t)}\boldsymbol{v}^{(t-1)}$ at time $t \in [T]$, with $\boldsymbol{v}^{(0)} = \boldsymbol{0}$. Applying (D.9) with radius R_τ, (D.8) bounds its coefficient by $5\pi\tau^2 \max_{|z|=R_\tau} |K(z)|^2$.

For $1 \le h \le k$, scalar filter algebra gives

$$\begin{aligned} \mathsf{K}_h^F\mathsf{D} &= \mathsf{F}_h, \qquad \mathsf{K}_h^F\mathsf{C} = \mathsf{K}_\star^F\mathsf{F}_{h-1}, \\ \mathsf{F}_h\boldsymbol{x}_j &= \overline{\boldsymbol{J}}_j\mathsf{K}_\star^F\mathsf{F}_{h-1}\boldsymbol{g}_j + \boldsymbol{c}_{j,h}, \qquad \boldsymbol{c}_{j,h} \coloneqq [\mathsf{K}_h^F, \overline{\boldsymbol{J}}_j]\boldsymbol{d}_j. \end{aligned} \tag{D.11}$$

The first identity on constantly extended $\boldsymbol{x}_j$ uses $F_h(1) = 0$. Applying (D.5) and (D.9) gives, whenever $0 < w_t \le 1$ and $w_t/x_j^{(t)}(a_j)$ has the required drift,

$$\sum_t w_t \left\| \boldsymbol{c}_{j,h}^{(t)} \right\|^2_{(\boldsymbol{x}_j^{(t)})^{-1}} \le 1000\kappa_1^2 \left(2\kappa_0 - 1\right)^2 \eta^2\tau^2\mathcal{P}_{j,T}. \tag{D.12}$$

Indeed $5\pi \cdot 60 < 1000$. If also $w_t \le \sqrt{\lambda_j^{(t)}}$, weighted convolution and (D.5) give

$$\sum_t w_t \left\| (\mathsf{F}_h \boldsymbol{x}_j)^{(t)} \right\|^2_{(\boldsymbol{x}_j^{(t)})^{-1}} \le 5\kappa_1^2 \mathcal{P}_{j,T}. \tag{D.13}$$

Payoff differences. For every $i \in [N]$ and $t \in \mathbb{Z}$, set $\boldsymbol{J}_{g,i}^{(t)} = \boldsymbol{J}_{\mathcal{G}_i}(\boldsymbol{x}^{(t)})$ and let $\boldsymbol{J}_{g,i}$ act pointwise on profile sequences. The centered payoff map $\mathcal{G}_i$ is multilinear in the entire profile. Each coordinate is the expectation of a pure-payoff difference in $[-1, 1]$ under the product distribution of actions. For every positive strategy profile $\boldsymbol{x}$ and tangents $\boldsymbol{z}, \boldsymbol{v}$ satisfying $\langle \mathbf{1}, \boldsymbol{z}_j \rangle = \langle \mathbf{1}, \boldsymbol{v}_j \rangle = 0$ for $j \in [N]$, independence gives

$$\begin{aligned} \|\boldsymbol{J}_{\mathcal{G}_i}(\boldsymbol{x})[\boldsymbol{z}]\|_\infty^2 \le & \sum_{j=1}^N \|\boldsymbol{z}_j\|^2_{\boldsymbol{x}_j^{-1}}, \\ \left\| \frac{\mathrm{d}}{\mathrm{d}s} \boldsymbol{J}_{\mathcal{G}_i}(\boldsymbol{x} + s\boldsymbol{v})[\boldsymbol{z}] \Big|_{s=0} \right\|_\infty^2 \le & 2 \left(\sum_{j=1}^N \|\boldsymbol{z}_j\|^2_{\boldsymbol{x}_j^{-1}} \right) \left(\sum_{j=1}^N \|\boldsymbol{v}_j\|^2_{\boldsymbol{x}_j^{-1}} \right). \end{aligned} \tag{D.14}$$

For the second inequality, sample independent actions $a_j \sim \boldsymbol{x}_j$, $j \in [N]$. The second derivative multiplies the payoff integrand by $\sum_{j \ne \ell} z_j(a_j) v_\ell(a_\ell) / (x_j(a_j) x_\ell(a_\ell))$. Different unordered pairs are orthogonal, and for $j \ne \ell$,

$$\mathbb{E}\left[\left(\frac{z_j(a_j) v_\ell(a_\ell) + v_j(a_j) z_\ell(a_\ell)}{x_j(a_j) x_\ell(a_\ell)} \right)^2 \right] \le 2 \left(\|\boldsymbol{z}_j\|^2_{\boldsymbol{x}_j^{-1}} \|\boldsymbol{v}_\ell\|^2_{\boldsymbol{x}_\ell^{-1}} + \|\boldsymbol{v}_j\|^2_{\boldsymbol{x}_j^{-1}} \|\boldsymbol{z}_\ell\|^2_{\boldsymbol{x}_\ell^{-1}} \right).$$

The Cauchy–Schwarz inequality proves (D.14).

For every $t \in [T]$, use $b_{j,t}$ from Lemma D.2. By (D.6), $\lambda_j^{(t)} \le 1$, and $e^{15\epsilon/4} \le 5/4$,

$$\sum_{j=1}^N b_{j,t}^4 \le \frac{5}{4} (2\kappa_0 - 1)^2 \eta^2 \sum_{j=1}^N P_j^{(t)}. \tag{D.15}$$

Fix $i \in [N]$ and a positive weight sequence $\boldsymbol{w} = (w_t)_{t=1}^T$ with $w_t \le 1$ such that every coordinate weight $w_t x_i^{(t)}(a_i)$ has logarithmic drift at most $1/(5\tau)$. Set $\|\boldsymbol{v}\|_{H_t}^2 = w_t \|\boldsymbol{v}\|^2_{\boldsymbol{x}_i^{(t)}}$. For $1 \le \ell < t$, integrate (D.14) along the profile segment from $\boldsymbol{x}^{(t-1)}$ to $\boldsymbol{x}^{(t)}$. Coordinate ratios in (D.5) and $\left\| (\mathsf{D}\boldsymbol{x}_j)^{(t)} \right\|^2_{(\boldsymbol{x}_j^{(t)})^{-1}} \le 5 b_{j,t}^2$ give

$$\begin{aligned} \left\| (\boldsymbol{J}_{g,i}^{(t)} - \boldsymbol{J}_{g,i}^{(t-1)})[(\mathsf{D}\boldsymbol{x})^{(\ell)}] \right\|_{H_t}^2 \le & 50 e^{2\epsilon(t-\ell+1)} \left(\sum_j b_{j,t}^2 \right) \left(\sum_j b_{j,\ell}^2 \right) \\ \le & 32 (2\kappa_0 - 1)^2 \eta^2 N e^{2\epsilon(t-\ell+1)} \sum_{j=1}^N \left(P_j^{(t)} + P_j^{(\ell)} \right). \end{aligned} \tag{D.16}$$

The second line uses the bound on the product of sums by $(N/2) \sum_j (b_{j,t}^4 + b_{j,\ell}^4)$ and $125/4 < 32$.

Parseval and the zero extension of $\mathsf{D}\boldsymbol{x}$ give

$$\begin{aligned} & \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\| (\boldsymbol{J}_{g,i}^{(t)} - \boldsymbol{J}_{g,i}^{(t-1)})[\left((R_\tau e^{\mathrm{i}\phi} I - \mathsf{L}_{\mathrm{in}})^{-1} \mathsf{D}\boldsymbol{x} \right)^{(t-1)}] \right\|_{H_t}^2 d\phi \\ = & \sum_{s \ge 0} R_\tau^{-2s-2} \left\| (\boldsymbol{J}_{g,i}^{(t)} - \boldsymbol{J}_{g,i}^{(t-1)})[(\mathsf{D}\boldsymbol{x})^{(t-1-s)}] \right\|_{H_t}^2. \end{aligned}$$

Since $2\epsilon \le 1/(8\tau)$,

$$\sum_{s \ge 0} R_\tau^{-2s-2} e^{2\epsilon(s+2)} = \frac{e^{4\epsilon}}{R_\tau^2 - e^{2\epsilon}} \le 5\tau.$$

Sum (D.16) over time in the preceding resolvent identity. Each of the two shifted movement sums is at most $\mathcal{P}_T$, so the integrated squared input to the middle operator of (D.10) is at most $320\,(2\kappa_0-1)^2\,\eta^2 N\tau\mathcal{P}_T$. Cauchy–Schwarz in (D.10) and (D.8) therefore give

$$\begin{aligned}\left\|[\mathsf{K}_h^F, \boldsymbol{J}_{g,i}]\mathsf{D}\boldsymbol{x}\right\|_{\ell^2(H)}^2 \leq& \frac{R_\tau^2\kappa_1^2}{R_\tau^2-\|\mathsf{L}_{\mathrm{out}}\|^2}320\,(2\kappa_0-1)^2\,\eta^2 N\tau\mathcal{P}_T\\ \leq& 1600\kappa_1^2\,(2\kappa_0-1)^2\,\eta^2 N\tau^2\mathcal{P}_T. \end{aligned}\tag{D.17}$$

Define $\boldsymbol{\zeta}_i^{(t)}=\mathbf{0}$ for $t\leq 1$ and, for $t\geq 2$, $\boldsymbol{\zeta}_i^{(t)}=\mathcal{G}_i(\boldsymbol{x}^{(t)})-\mathcal{G}_i(\boldsymbol{x}^{(t-1)})-\boldsymbol{J}_{g,i}^{(t)}[(\mathsf{D}\boldsymbol{x})^{(t)}]$. The integral remainder in (D.14) has weight integrating to $1/2$, so $\left\|\boldsymbol{\zeta}_i^{(t)}\right\|_\infty \leq 4\sum_j b_{j,t}^2$; the coefficient before rounding is $5e^{2\epsilon}/\sqrt{2}<4$. By (D.15) and weighted convolution,

$$\begin{aligned}\sum_t \left\|\boldsymbol{\zeta}_i^{(t)}\right\|_{H_t}^2 \leq& 20\,(2\kappa_0-1)^2\,\eta^2 N\mathcal{P}_T,\\ \left\|\mathsf{K}_h^F\boldsymbol{\zeta}_i\right\|_{\ell^2(H)}^2 \leq& 20\kappa_1^2\,(2\kappa_0-1)^2\,\eta^2 N\mathcal{P}_T.\end{aligned}$$

The initialization gives $\mathsf{D}\boldsymbol{g}_i=\boldsymbol{J}_{g,i}\mathsf{D}\boldsymbol{x}+\boldsymbol{\zeta}_i+(\boldsymbol{g}_i^{(1)}\,\mathbb{1}_{\{t=1\}})_{t\in\mathbb{Z}}$. The last term after applying K_h^F has squared weighted norm at most κ_1^2, by Parseval on the unit circle and $\left\|\boldsymbol{g}_i^{(1)}\right\|_\infty\leq 1$. Combining (D.17) with these two estimates gives, for every $1\leq h\leq k$,

$$\begin{aligned}\mathsf{F}_h\boldsymbol{g}_i =& \boldsymbol{J}_{g,i}\mathsf{F}_h\boldsymbol{x}+\boldsymbol{\chi}_{i,h},\\ \boldsymbol{\chi}_{i,h} :=& [\mathsf{K}_h^F,\boldsymbol{J}_{g,i}]\mathsf{D}\boldsymbol{x}+\mathsf{K}_h^F\boldsymbol{\zeta}_i+\mathsf{K}_h^F\left((\boldsymbol{g}_i^{(1)}\underset{\{t=1\}}{\mathbb{1}})_{t\in\mathbb{Z}}\right),\\ \sum_t w_t\left\|\boldsymbol{\chi}_{i,h}^{(t)}\right\|_{\boldsymbol{x}_i^{(t)}}^2 \leq& 3\kappa_1^2+5000\kappa_1^2\,(2\kappa_0-1)^2\,\eta^2 N\tau^2\mathcal{P}_T.\end{aligned}\tag{D.18}$$

Here $3(1600N\tau^2+20N)\leq 5000N\tau^2$, since $\tau\geq 2$.

The finite-player recursion. For $r\in[N]$ and $S\subseteq[N]\setminus\{r\}$, define the weight sequence $\widetilde{\boldsymbol{w}}_{r,S}=(\widetilde{w}_{r,S}^{(t)})_{t\in\mathbb{Z}}$ by

$$\widetilde{w}_{r,S}^{(t)} := \min\left(\left\{\left(\lambda_r^{(t)}\right)^{3/2}\right\}\cup\left\{\sqrt{\lambda_j^{(t)}} : j\in S\right\}\right).$$

Their logarithmic drift is at most $3\epsilon/2$. All coordinate weights $\widetilde{\boldsymbol{w}}_{r,S}\boldsymbol{x}_j$, $\widetilde{\boldsymbol{w}}_{r,S}/\boldsymbol{x}_j$, and $\sqrt{\boldsymbol{\lambda}_j}/\boldsymbol{x}_j$ consequently have drift at most $7\epsilon/2\leq 1/(5\tau)$. For $j\in[N]$, $1\leq h\leq k$, and a positive weight sequence $\boldsymbol{w}=(w_t)_{t=1}^T$, define

$$\mathcal{E}_j^h(\boldsymbol{w}) := \sum_{t=1}^T w_t\left\|(\mathsf{F}_h\boldsymbol{x}_j)^{(t)}\right\|_{(\boldsymbol{x}_j^{(t)})^{-1}}^2.$$

If $j\in S\cup\{r\}$, (D.13) gives $\mathcal{E}_j^h(\widetilde{\boldsymbol{w}}_{r,S})\leq 5\kappa_1^2\mathcal{P}_{j,T}$.

Suppose $j\notin S\cup\{r\}$ and $|S|\leq k-2$. Set $h=k-|S|\geq 2$, $\boldsymbol{w}=\widetilde{\boldsymbol{w}}_{r,S}$, and $\boldsymbol{w}'=\widetilde{\boldsymbol{w}}_{r,S\cup\{j\}}$. Split $\boldsymbol{w}=\boldsymbol{w}_{\mathrm{hi}}+\boldsymbol{w}_{\mathrm{lo}}$ pointwise by $\boldsymbol{w}_{\mathrm{hi}}=\min(\boldsymbol{w},N\sqrt{\boldsymbol{\lambda}_j})$. The direct estimate is at most $5\kappa_1^2 N\mathcal{P}_{j,T}$ on $\boldsymbol{w}_{\mathrm{hi}}$, since it is bounded above by $N\sqrt{\boldsymbol{\lambda}_j}$. The other part satisfies

$$\boldsymbol{w}_{\mathrm{lo}}\boldsymbol{\lambda}_j\leq\frac{\boldsymbol{w}'}{4N}.\tag{D.19}$$

Indeed, for every $t\in[T]$ with $w_{\mathrm{lo}}^{(t)}>0$, we have $\sqrt{\lambda_j^{(t)}}<w_t/N$ and $w_t'=\sqrt{\lambda_j^{(t)}}$. Thus

$$\frac{w_{\mathrm{lo}}^{(t)}\lambda_j^{(t)}}{w_t'}=(w_t-N\sqrt{\lambda_j^{(t)}})\sqrt{\lambda_j^{(t)}}\leq\frac{w_t^2}{4N}\leq\frac{1}{4N}.$$

No convolution estimate is applied with $\boldsymbol{w}_{\mathrm{hi}}$ or $\boldsymbol{w}_{\mathrm{lo}}$. For the commutator on $\boldsymbol{w}_{\mathrm{lo}}$, use $\boldsymbol{w}_{\mathrm{lo}} \le 1$ and apply (D.12) with weight 1; the coordinate weights $1/x_j^{(t)}(a_j)$ have drift at most 2ϵ.

Apply (D.5), (D.11), (D.12) and (D.19) and then convolve in the weight $\boldsymbol{w}'\boldsymbol{x}_j$. This gives

$$\begin{aligned}\mathcal{E}_j^h(\boldsymbol{w}) \le& \kappa_1^2 \left(5N + 2000\left(2\kappa_0 - 1\right)^2 \eta^2\tau^2\right) \mathcal{P}_{j,T} \\ &+ \frac{5\eta^2\tau^2\kappa_2^2}{2N} \sum_t w_t' \left\| \left(\mathsf{F}_{h-1}\boldsymbol{g}_j\right)^{(t)} \right\|_{\boldsymbol{x}_j^{(t)}}^2 .\end{aligned}$$

The coefficient in the last line is $2 \cdot 5\eta^2 \cdot (4N)^{-1} \cdot (\tau\kappa_2)^2$. By (D.14) and (D.18),

$$\begin{aligned}\sum_t w_t' \left\| \left(\mathsf{F}_{h-1}\boldsymbol{g}_j\right)^{(t)} \right\|_{\boldsymbol{x}_j^{(t)}}^2 \le& 2\sum_\ell \mathcal{E}_\ell^{h-1}(\boldsymbol{w}') + 6\kappa_1^2 \\ &+ 10000\kappa_1^2\left(2\kappa_0 - 1\right)^2 \eta^2 N\tau^2\mathcal{P}_T. \end{aligned} \tag{D.20}$$

When $|S| = k-1$, the identity $\mathsf{F}_1\boldsymbol{x}_j = \mathsf{K}_1^F \mathsf{D}\boldsymbol{x}_j$, weighted convolution in $\widetilde{\boldsymbol{w}}_{r,S}/\boldsymbol{x}_j$, and (D.5) give, for every $j \notin S \cup \{r\}$,

$$\begin{aligned}\mathcal{E}_j^1(\widetilde{\boldsymbol{w}}_{r,S}) \le& \kappa_1^2 \sum_t \widetilde{w}_{r,S}^{(t)} \left\| \left(\mathsf{D}\boldsymbol{x}_j\right)^{(t)} \right\|_{(\boldsymbol{x}_j^{(t)})^{-1}}^2 \\ \le& 5\kappa_1^2\eta^2 \sum_t \widetilde{w}_{r,S}^{(t)} \lambda_j^{(t)} \left\| \boldsymbol{d}_j^{(t)} \right\|_{\boldsymbol{x}_j^{(t)}}^2 \le 5\kappa_1^2\left(2\kappa_0 - 1\right)^2\eta^2 T.\end{aligned}$$

There are $N - k$ such players. Use q from (D.1). For $0 \le s \le k-1$, define

$$\mathcal{E}_s \coloneqq \max_{r,S:\,|S|=s} \sum_j \mathcal{E}_j^{k-s}(\widetilde{\boldsymbol{w}}_{r,S}),$$

where the maximum is over $r \in [N]$ and $S \subseteq [N] \setminus \{r\}$ of the stated size. Summing the two membership cases and using (D.20) gives the following bounds, the first for $0 \le s < k-1$:

$$\begin{aligned}\mathcal{E}_s \le& q\mathcal{E}_{s+1} + 3\kappa_1^2 q \\ &+ \kappa_1^2\left[5N + \left(2000 + 5000Nq\right)\left(2\kappa_0 - 1\right)^2\eta^2\tau^2\right]\mathcal{P}_T, \\ \mathcal{E}_{k-1} \le& 5\kappa_1^2\mathcal{P}_T + 5\kappa_1^2\left(2\kappa_0 - 1\right)^2\eta^2(N-k)T.\end{aligned} \tag{D.21}$$

Since $q < 1$, backward induction in (D.21) gives

$$\begin{aligned}\mathcal{E}_0 \le& \frac{\kappa_1^2}{1-q}\left\{3q + \left[5N + \left(2000 + 5000Nq\right)\left(2\kappa_0 - 1\right)^2\eta^2\tau^2\right]\mathcal{P}_T\right\} \\ &+ 5\kappa_1^2\left(2\kappa_0 - 1\right)^2\eta^2(N-k)q^{k-1}T.\end{aligned}$$

This includes $k = 1$, for which no recursive substitution is needed. Apply (D.14) and (D.18) at $h = k$ with $\boldsymbol{w} = (\boldsymbol{\lambda}_i)^{3/2}$ and use $\mathsf{F}_k = \mathsf{F}$. The resulting estimate is

$$\begin{aligned}\sum_t (\lambda_i^{(t)})^{3/2} \left\| \left(\mathsf{F}\boldsymbol{g}_i\right)^{(t)} \right\|_{\boldsymbol{x}_i^{(t)}}^2 \le& \frac{6\kappa_1^2}{1-q} \\ &+ \frac{\kappa_1^2}{1-q}\left[10N + \left(4000 + 10000N\right)\left(2\kappa_0 - 1\right)^2\eta^2\tau^2\right]\mathcal{P}_T \\ &+ 10\kappa_1^2\left(2\kappa_0 - 1\right)^2\eta^2(N-k)q^{k-1}T \\ =& \alpha_{0,T} + \alpha_1\mathcal{P}_T.\end{aligned} \tag{D.22}$$

Insert (D.22) into (D.7). Summing the nonnegative lifted regrets gives

$$\mathcal{P}_T \le \frac{\sum_j r_j + 3\eta^2 N\alpha_{0,T}}{1/2 - \beta}.$$

Substitution into the individual inequality gives

$$\mathcal{R}_{i,T} \le \frac{1}{\eta}\left(r_i + \frac{\beta}{1/2-\beta}\frac{1}{N}\sum_j r_j\right) + \frac{3\eta\alpha_{0,T}}{1-2\beta}.$$

The comparator bound proves (D.2).

Finally suppose $0 < \eta \le (1000\tau(2\kappa_0-1)\kappa_1\kappa_2)^{-1}$. Then $\epsilon \le 1/(1000\tau)$, $q \le 5/10^6$, and $(2\kappa_0-1)^2\eta^2\tau^2 \le 10^{-6}$. Thus $\alpha_0 \le 7\kappa_1^2$ and $\alpha_1 \le 11N\kappa_1^2$; the latter follows from $(10.01N + 0.004)/(1-5\cdot10^{-6}) < 11N$ for $N \ge 2$. Consequently $\beta \le 33/10^6$, and

$$\begin{aligned}\mathcal{R}_{i,T} \le& \frac{1}{\eta}\left(r_i + \frac{1}{N}\sum_j r_j\right) + 22\eta\kappa_1^2 + \frac{3\eta(\alpha_{0,T}-\alpha_0)}{1-2\beta}\\ \le& \frac{2}{\eta}\left(r_i + \frac{1}{N}\sum_j r_j\right) + 60\kappa_1^2(2\kappa_0-1)^2(N-k)\eta^3 q^{k-1}T.\end{aligned}$$

The last step uses $\eta^2\kappa_1^2 \le 1/(10^6\tau^2)$, $r_i \ge 3/2$, and $(1-2\beta)^{-1} < 2$. Multiplication by M_i proves (D.3). □

Proof of Theorem 5.2. Fix F, N and let k be as in Section 5. By Lemma 4.1, $M_i \le e$ and $r_i \le e(1+\log m_i)$ for every $i \in [N]$. If $k = 0$, (D.24) gives the claimed bound for $0 < \eta \le 1/(20\kappa_0)$. If $k \ge 1$, choose $\tau, \kappa_0, \kappa_1, \kappa_2$ as in Section D.1. For every $0 < \eta \le (1000\tau(2\kappa_0-1)\kappa_1\kappa_2)^{-1}$, (D.3) applies. Substituting $q = 5\eta^2\tau^2\kappa_2^2$ and using $m_i \ge 2$ gives the claimed bound, since k, τ and the three filter norms depend only on F, N. If $k = N$, the horizon term vanishes. □

D.2 ECHO-OFTRL SPECIALIZATION

Theorem 4.2. *In every N-player finite normal-form game with utilities in $[0,1]$, suppose every player runs Algorithm 1 with the common learning rate $\eta = 1/(2000N)$. Recall that $r_i = (p_i+1)m_i^{1/p_i}/2$ for every $i \in [N]$. Then, for every $i \in [N]$ and horizon $T \ge 1$,*

$$[\mathrm{Reg}_i(T)]_+ \le m_i^{1/p_i}\left(2000Nr_i + \sum_{j=1}^N r_j + \frac{3}{40N}\right) \le O\left(N\log m_{\max}\right).$$

Proof. Recall that $\delta(z) = (1-z)/(1-\rho z)$ is the transfer function of δ in (4.4). By (4.5), take $F(z) = (1-z)\delta(z)^{N-1}$ and $k = \tau = N$. The definitions in Sections D.1 and 5.3 give $\mathsf{F}_0 = I - \rho\mathsf{L}$ and, for $1 \le h \le N$,

$$\mathsf{F}_h = \mathsf{D}\delta^{h-1}, \qquad \mathsf{K}_h^F = \delta^{h-1}, \qquad \mathsf{K}_\star^F = \mathsf{A} - \mathsf{D}\delta^N.$$

We first bound the absolute coefficient sums.

Lemma D.5. *For every $N \ge 2$, set $\rho = 1 - 1/N$ and $\delta(z) = (1-z)/(1-\rho z)$. Then*

$$\left\|1 - (1-z)\delta(z)^{N-1}\right\|_{\ell^1} \le 48.$$

Its proof is in Section E. By Lemma D.5, $\kappa_0 = 1 + \|I - \mathsf{F}\|_{\ell^1} \le 49$, so $2\kappa_0 - 1 \le 97$. On $|z| = R_\tau$, $\rho R_\tau^2 < 1$ implies that $|\delta(z)|$ is maximized at $z = -R_\tau$. Since $R_\tau \le 1 + 1/(4N)$,

$$\begin{aligned}\max_{|z|=R_\tau}|\delta(z)| \le& 1 + \frac{1}{N}, \qquad \max_{|z|=R_\tau}\left|\frac{1}{1-\rho z}\right| \le \frac{4N}{3},\\ \max_{|z|=R_\tau}\left|(1-z)\delta(z)^N\right| \le& \frac{9}{4}e \le \frac{297}{32}.\end{aligned}$$

Thus $\kappa_1 \le e$, so $\kappa_1^2 < 8$, and $\kappa_2 \le 4/3 + 297/64 < 6$. At $\eta = 1/(2000N)$, $\eta(2\kappa_0-1) \le 97/(2000N) < 1/(20N)$, $q \le 0.000045$, $\alpha_0 < 49$, $\alpha_1 < 306N$, and $\beta < 0.0002295$. Substitution into (D.2) proves Theorem 4.2, since $M_i = m_i^{1/p_i}$ and

$$\frac{2000\beta}{1/2-\beta} < 1, \qquad \frac{147}{2000(1-2\beta)} < \frac{3}{40}.$$

□

D.3 THE ONE-STEP PREDICTOR

Corollary D.6. *Take $F(z) = 1 - z$, $k = 1$, and $\tau = N$. Then $\kappa_0 = 2$, $\kappa_1 = 1$, and $\kappa_2 \le 5$. For every $0 < \eta \le 1/(15000N)$, every $i \in [N]$, and every $T \ge 1$,*

$$[\mathrm{Reg}_i(T)]_+ \le \frac{2M_i}{\eta}\left(r_i + \frac{1}{N}\sum_j r_j\right) + 540M_i(N-1)\eta^3 T. \tag{D.23}$$

For a specified horizon $T \ge 1$, choose the common rate

$$\eta = \min\left\{\frac{1}{15000N}, \left(\frac{1+\log m_{\max}}{NT}\right)^{1/4}\right\}.$$

Each player's positive regret at horizon T is at most $O(N\log m_{\max} + N^{1/4}(\log m_{\max})^{3/4}T^{1/4})$, with a universal implicit constant.

Proof. Here $\mathsf{K}_1^F = I$ and $\mathsf{C} = 2\mathsf{L} - \mathsf{L}^2$. On $|z| = R_\tau$, the bounds $R_\tau \le 9/8$ and $1/|1-\rho_\tau z| \le 4N/3$ give

$$\frac{1}{N}\left|K_\star^F(z)\right| \le \frac{4}{3}\left(2R_\tau + R_\tau^2\right) \le \frac{75}{16} < 5.$$

Thus (D.3) gives (D.23). Since $M_i \le e$ and $r_i \le e(1+\log m_{\max})$, its right side is at most $O((1+\log m_{\max})/\eta + N\eta^3 T)$. If the minimum defining η is its second argument, both terms have order $N^{1/4}(1+\log m_{\max})^{3/4}T^{1/4}$. Otherwise $N\eta^4 T \le 1 + \log m_{\max}$, and both terms are at most $O(N(1+\log m_{\max}))$. □

D.4 ADVERSARIAL REGRET FOR GENERAL FILTERS

The same optimistic-FTRL estimate applies even without a zero at 1: for every real causal filter $F(z) = \sum_{s\ge 0} f_s z^s$ with $f_0 = 1$ and $\kappa_0 = \sum_{s\ge 0}|f_s| < \infty$, the update in (5.1), with $0 < \eta \le 1/(20\kappa_0)$, satisfies

$$[\mathrm{Reg}_i(T)]_+ \le M_i\left(\frac{r_i}{\eta} + 3\eta T\kappa_0^2\right) \qquad (i \in [N],\ T \ge 1). \tag{D.24}$$

Indeed, (D.7) requires only the local prediction-error condition $2\eta\left\|\boldsymbol{e}_i^{(t)}\right\|_\infty \le 1/10$, independently of the zero multiplicity. Using $\lambda_i^{(t)} \le 1$ and $\left\|\boldsymbol{e}_i^{(t)}\right\|_\infty \le \kappa_0$ proves (D.24). For fixed F and action counts, choosing $\eta = \Theta(T^{-1/2})$ subject to the preceding upper bound gives adversarial regret at most $O(\sqrt{T})$. The transfer-function estimates in Theorem D.1 reduce the horizon term, and remove it when $k = N$, through the same negative movement term in (D.7).

E RESPONSE AND FILTER ESTIMATES

Lemma D.2. *Let $0 < \epsilon \le 1/25$. For every $i \in [N]$, consider scores $\boldsymbol{\theta}_i^{(t)}$ satisfying $\left\|\boldsymbol{\theta}_i^{(t)} - \boldsymbol{\theta}_i^{(t-1)}\right\|_\infty \le \epsilon$ for $t \in [T]$, with constant extension to $t \le 0$. Use the responses and quantities in* (D.4)*. For every $t \in [T]$, $a_i \in [m_i]$, and $\boldsymbol{v} \in \mathbb{R}^{m_i}$,*

$$\begin{aligned}
\left|\log\frac{\lambda_i^{(t)}}{\lambda_i^{(t-1)}}\right| &\le \epsilon, \qquad \left|\log\frac{x_i^{(t)}(a_i)}{x_i^{(t-1)}(a_i)}\right| \le 2\epsilon,\\
\sqrt{\lambda_i^{(t)}}\left\|(\mathsf{D}\boldsymbol{x}_i)^{(t)}\right\|^2_{(\boldsymbol{x}_i^{(t)})^{-1}} &\le 5P_i^{(t)},\\
\left\|\overline{\boldsymbol{J}}_i^{(t)}\boldsymbol{v}\right\|^2_{(\boldsymbol{x}_i^{(t)})^{-1}} &\le 5\eta^2\lambda_i^{(t)}\left\|\boldsymbol{v}\right\|^2_{\boldsymbol{x}_i^{(t)}},
\end{aligned} \tag{D.5}$$

$$\sum_{t=1}^{T} \left\| \overline{\boldsymbol{J}}_i^{(t)} - \overline{\boldsymbol{J}}_i^{(t-1)} \right\|^2_{\infty\to(\boldsymbol{x}_i^{(t)})^{-1}} \le 60\eta^2 \mathcal{P}_{i,T}.$$

For every $t \in [T]$, set $b_{i,t} \coloneqq \left\| \sqrt{\boldsymbol{x}_i^{(t)}} - \sqrt{\boldsymbol{x}_i^{(t-1)}} \right\|_2$, where square roots are coordinatewise. Then

$$b_{i,t} \le \epsilon e^{\epsilon/2} \sqrt{\lambda_i^{(t)}}, \qquad \sqrt{\lambda_i^{(t)}} b_{i,t}^2 \le e^{11\epsilon/4} P_i^{(t)}. \tag{D.6}$$

Proof. Fix a player $i \in [N]$ and suppress its index. Set $m = m_i$, $p = p_i$, $M = M_i$, $r = r_i$, and $\psi = \psi_i$, and let ψ^* denote its Fenchel conjugate. Recall $h(s) = (s + \sqrt{1+s^2})/2$ for $s \in \mathbb{R}$ from the proof of Lemma 4.1. All coordinate indices range over $[m]$. Expectations and variances are over a coordinate drawn from the indicated probability vector. For an unscaled score $\boldsymbol{s} \in \mathbb{R}^m$, define the vectors $\boldsymbol{h}, \boldsymbol{\gamma}, \boldsymbol{\pi} \in \mathbb{R}^m$ and the unscaled responses by

$$\begin{aligned} h_a =& h(s(a)), \qquad h'_a = h'(s(a)), \\ y^{\mathrm{u}}(a) =& \left(h_a / \left\|\boldsymbol{h}\right\|_p \right)^{p-1} h'_a, \qquad \lambda^{\mathrm{u}} = \sum_a y^{\mathrm{u}}(a), \qquad \boldsymbol{x}^{\mathrm{u}} = \boldsymbol{y}^{\mathrm{u}} / \lambda^{\mathrm{u}}, \\ \gamma(a) =& (p + 1 - 2h'_a) \frac{h'_a}{h_a}, \qquad \pi(a) = h_a^p / \left\|\boldsymbol{h}\right\|_p^p. \end{aligned}$$

Here primes denote derivatives with respect to the corresponding score coordinate, and the superscript u denotes unscaled quantities. Regard these quantities as functions of $\boldsymbol{s}$. Thus $\boldsymbol{y}^{\mathrm{u}} = \nabla_{\boldsymbol{s}} \left\|\boldsymbol{h}\right\|_p$. Define $\boldsymbol{H}^{\mathrm{u}}(\boldsymbol{s}) = \nabla^2_{\boldsymbol{s}} \left\|\boldsymbol{h}\right\|_p$ and $\boldsymbol{J}^{\mathrm{u}}(\boldsymbol{s}) = \boldsymbol{J}_{\boldsymbol{x}^{\mathrm{u}}}(\boldsymbol{s})$, suppressing their arguments when unambiguous. Products of coordinate vectors below are coordinatewise. Direct differentiation gives

$$\begin{aligned} \boldsymbol{H}^{\mathrm{u}} =& \operatorname{diag}(\boldsymbol{y}^{\mathrm{u}} \boldsymbol{\gamma}) - \frac{p-1}{\left\|\boldsymbol{h}\right\|_p} \boldsymbol{y}^{\mathrm{u}} \left(\boldsymbol{y}^{\mathrm{u}}\right)^\top, \\ \boldsymbol{J}^{\mathrm{u}} =& \left(\operatorname{diag}(\boldsymbol{x}^{\mathrm{u}}) - \boldsymbol{x}^{\mathrm{u}} \left(\boldsymbol{x}^{\mathrm{u}}\right)^\top \right) \operatorname{diag}(\boldsymbol{\gamma}). \end{aligned} \tag{E.1}$$

For every $\boldsymbol{v} \in \mathbb{R}^m$, the Hessian has the positive decomposition

$$\begin{aligned} \boldsymbol{v}^\top \boldsymbol{H}^{\mathrm{u}} \boldsymbol{v} =& E_1 + E_2, \\ E_1 =& \sum_a \left(h_a / \left\|\boldsymbol{h}\right\|_p \right)^{p-1} h''(s(a)) v(a)^2, \\ E_2 =& (p-1) \left\|\boldsymbol{h}\right\|_p \operatorname{Var}_{\boldsymbol{\pi}} \left(\frac{h'_a}{h_a} v(a) \right). \end{aligned} \tag{E.2}$$

In particular the Hessian is positive definite. The elementary identities

$$\begin{aligned} h'_a =& \frac{4h_a^2}{1+4h_a^2}, \qquad \left(\frac{h'_a}{h_a}\right)^2 = 4h'_a(1-h'_a), \\ \frac{h''(s(a))}{h'_a} =& 2\frac{h'_a}{h_a}(1-h'_a), \\ \lambda^{\mathrm{u}} \ge& \frac{4\left\|\boldsymbol{h}\right\|_p^2}{M(1+4\left\|\boldsymbol{h}\right\|_p^2)}, \qquad 0 < \lambda^{\mathrm{u}} < M, \qquad h'_a \le M\lambda^{\mathrm{u}} \end{aligned}$$

imply

$$\gamma_{\max} \coloneqq \max_a \gamma(a) \le p+1, \qquad \gamma_{\max}^2 \le 4(p+1)^2 M \lambda^{\mathrm{u}}. \tag{E.3}$$

For the lower bound on λ^{u}, use $h_a \le \left\|\boldsymbol{h}\right\|_p$ and $\sum_a h_a^{p+1} \ge \left\|\boldsymbol{h}\right\|_p^{p+1} / M$. These also give $h'_a \le 4\left\|\boldsymbol{h}\right\|_p^2 / (1 + 4\left\|\boldsymbol{h}\right\|_p^2) \le M\lambda^{\mathrm{u}}$.

Set $z(a) = \gamma(a) v(a)$. Testing positive semidefiniteness of $\boldsymbol{H}^{\mathrm{u}}$ on $(1/\gamma(a))_{a\in[m]}$ gives

$$\frac{(p-1)\lambda^{\mathrm{u}}}{\left\|\boldsymbol{h}\right\|_p} \mathbb{E}_{\boldsymbol{x}^{\mathrm{u}}} \left[1/\gamma\right] \le 1.$$

Consequently,

$$\begin{aligned}\frac{\boldsymbol{v}^\top \boldsymbol{H}^{\mathrm{u}} \boldsymbol{v}}{\lambda^{\mathrm{u}}} =& \mathbb{E}_{\boldsymbol{x}^{\mathrm{u}}}\left[z^2/\gamma\right] - \frac{(p-1)\lambda^{\mathrm{u}}}{\|\boldsymbol{h}\|_p}\left(\mathbb{E}_{\boldsymbol{x}^{\mathrm{u}}}\left[z/\gamma\right]\right)^2 \\ \geq& \min_{b\in\mathbb{R}} \mathbb{E}_{\boldsymbol{x}^{\mathrm{u}}}\left[(z-b)^2/\gamma\right] \geq \frac{\mathrm{Var}_{\boldsymbol{x}^{\mathrm{u}}}(z)}{\gamma_{\max}}. \end{aligned} \tag{E.4}$$

Also

$$\begin{aligned}\|\boldsymbol{J}^{\mathrm{u}}\boldsymbol{v}\|^2_{(\boldsymbol{x}^{\mathrm{u}})^{-1}} =& \mathrm{Var}_{\boldsymbol{x}^{\mathrm{u}}}(\boldsymbol{\gamma}\boldsymbol{v}), \qquad \left\|\boldsymbol{J}_{\sqrt{\boldsymbol{x}^{\mathrm{u}}}}(\boldsymbol{s})[\boldsymbol{v}]\right\|_2^2 = \frac{1}{4}\mathrm{Var}_{\boldsymbol{x}^{\mathrm{u}}}(\boldsymbol{\gamma}\boldsymbol{v}), \\ \langle \nabla_{\boldsymbol{s}} \log \lambda^{\mathrm{u}}, \boldsymbol{v}\rangle =& \mathbb{E}_{\boldsymbol{x}^{\mathrm{u}}}\left[\left(\gamma(a) - \frac{(p-1)\lambda^{\mathrm{u}}}{\|\boldsymbol{h}\|_p}\right) v(a)\right], \\ \langle \nabla_{\boldsymbol{s}} \log x^{\mathrm{u}}(a), \boldsymbol{v}\rangle =& \gamma(a)v(a) - \mathbb{E}_{\boldsymbol{x}^{\mathrm{u}}}\left[\boldsymbol{\gamma}\boldsymbol{v}\right].\end{aligned}$$

Here $\lambda^{\mathrm{u}}/\|\boldsymbol{h}\|_p = \mathbb{E}_{\boldsymbol{\pi}}\left[h_a'/h_a\right] \leq 1$, so the two terms inside the mass coefficient lie in $[0, p+1]$ and $[0, p-1]$. For $\boldsymbol{\theta} \in \mathbb{R}^m$, write $\boldsymbol{H}(\boldsymbol{\theta}) = \nabla^2\psi^*(\boldsymbol{\theta})$ and $\boldsymbol{J}(\boldsymbol{\theta}) = \boldsymbol{J}_{\boldsymbol{x}}(\boldsymbol{\theta})$, and suppress their score arguments when unambiguous. At $\boldsymbol{s} = \boldsymbol{\theta}/(2r)$ the scaling identities are

$$\begin{aligned}\boldsymbol{y} =& \boldsymbol{y}^{\mathrm{u}}/M, \qquad \lambda = \lambda^{\mathrm{u}}/M, \qquad \boldsymbol{x} = \boldsymbol{x}^{\mathrm{u}}, \\ \boldsymbol{H} =& \boldsymbol{H}^{\mathrm{u}}/(2rM), \qquad \boldsymbol{J} = \boldsymbol{J}^{\mathrm{u}}/(2r).\end{aligned}$$

Use $\preceq$ for the positive-semidefinite order on symmetric matrices. Since $2r = (p+1)M$, the preceding formulas prove

$$\begin{aligned}\boldsymbol{H} \preceq& 2\lambda^{3/2}\,\mathrm{diag}(\boldsymbol{x}), \\ \|\boldsymbol{J}\boldsymbol{v}\|^2_{\boldsymbol{x}^{-1}} \leq& 4\lambda\|\boldsymbol{v}\|^2_{\boldsymbol{x}}, \qquad \left\|\boldsymbol{J}_{\sqrt{\boldsymbol{x}}}(\boldsymbol{\theta})[\boldsymbol{v}]\right\|_2 \leq \sqrt{\lambda}\,\|\boldsymbol{v}\|_\infty, \\ \sqrt{\lambda}\left\|\boldsymbol{J}_{\sqrt{\boldsymbol{x}}}(\boldsymbol{\theta})[\boldsymbol{v}]\right\|_2^2 \leq& \tfrac{1}{2}\boldsymbol{v}^\top\boldsymbol{H}\boldsymbol{v}, \\ |\langle \nabla \log\lambda(\boldsymbol{\theta}), \boldsymbol{v}\rangle| \leq& \|\boldsymbol{v}\|_\infty/M, \qquad |\langle \nabla_{\boldsymbol{\theta}} \log x(\boldsymbol{\theta})(a), \boldsymbol{v}\rangle| \leq 2\|\boldsymbol{v}\|_\infty/M.\end{aligned} \tag{E.5}$$

Fix $\boldsymbol{v}, \boldsymbol{\xi} \in \mathbb{R}^m$. Differentiating (E.2) in direction $\boldsymbol{\xi}$, with $\boldsymbol{v}$ fixed, gives

$$\begin{aligned}|\langle \nabla_{\boldsymbol{s}} E_1, \boldsymbol{\xi}\rangle| \leq& (2p - \tfrac{1}{2})\|\boldsymbol{\xi}\|_\infty E_1, \\ |\langle \nabla_{\boldsymbol{s}} E_2, \boldsymbol{\xi}\rangle| \leq& \left[(2p+1)E_2 + 2\sqrt{2(p-1)E_1E_2}\right]\|\boldsymbol{\xi}\|_\infty.\end{aligned}$$

Indeed, $\left|\left\langle \nabla_{\boldsymbol{s}} \log\|\boldsymbol{h}\|_p, \boldsymbol{\xi}\right\rangle\right| \leq \|\boldsymbol{\xi}\|_\infty$, $|\langle \nabla_{\boldsymbol{s}} \log\pi(a), \boldsymbol{\xi}\rangle| \leq 2p\|\boldsymbol{\xi}\|_\infty$, and $|\mathrm{d}\log h''(s)/\mathrm{d}s| \leq 3/2$. For the mixed term, set $z(a) = (h_a'/h_a)v(a)$. Differentiating the variance explicitly yields

$$\begin{aligned}\langle \nabla_{\boldsymbol{s}} \mathrm{Var}_{\boldsymbol{\pi}}(\boldsymbol{z}), \boldsymbol{\xi}\rangle =& \sum_a \langle \nabla_{\boldsymbol{s}}\pi(a), \boldsymbol{\xi}\rangle \left(z(a) - \mathbb{E}_{\boldsymbol{\pi}}\left[z\right]\right)^2 + 2\mathbb{E}_{\boldsymbol{\pi}}\left[\left(z - \mathbb{E}_{\boldsymbol{\pi}}\left[z\right]\right)\langle \nabla_{\boldsymbol{s}} z, \boldsymbol{\xi}\rangle\right], \\ \mathbb{E}_{\boldsymbol{\pi}}\left[\langle \nabla_{\boldsymbol{s}} z, \boldsymbol{\xi}\rangle^2\right] \leq& 4\|\boldsymbol{\xi}\|^2_\infty \mathbb{E}_{\boldsymbol{\pi}}\left[z^2(1-h_a')\right] = \frac{2E_1}{\|\boldsymbol{h}\|_p}\|\boldsymbol{\xi}\|^2_\infty.\end{aligned}$$

The larger eigenvalue of the resulting quadratic form in $(\sqrt{E_1}, \sqrt{E_2})$ satisfies

$$2p + \frac{1}{4} + \sqrt{2p - \frac{23}{16}} \leq \frac{9(p+1)}{4}.$$

Squaring reduces this inequality to $(p-8)^2 + 23 \geq 0$. Scaling therefore gives

$$\left|\frac{\mathrm{d}}{\mathrm{d}a}\left[\boldsymbol{v}^\top \boldsymbol{H}(\boldsymbol{\theta} + a\boldsymbol{\xi})\boldsymbol{v}\right]\Big|_{a=0}\right| \leq \frac{9}{4}\|\boldsymbol{\xi}\|_\infty \boldsymbol{v}^\top\boldsymbol{H}\boldsymbol{v}. \tag{E.6}$$

For $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^m$ and $a \in [m]$, differentiating (E.1) gives

$$\frac{\frac{\mathrm{d}}{\mathrm{d}q}\left(\boldsymbol{J}^{\mathrm{u}}(\boldsymbol{s} + q\boldsymbol{v})\boldsymbol{w}\right)(a)\Big|_{q=0}}{x^{\mathrm{u}}(a)} = \left(\gamma(a)v(a) - \mathbb{E}_{\boldsymbol{x}^{\mathrm{u}}}\left[\gamma v\right]\right)\left(\gamma(a)w(a) - \mathbb{E}_{\boldsymbol{x}^{\mathrm{u}}}\left[\gamma w\right]\right)$$

$$- \operatorname{Cov}_{\boldsymbol{x}^{\mathrm{u}}}(\gamma v, \gamma w) + \gamma'(a) v(a) w(a) - \mathbb{E}_{\boldsymbol{x}^{\mathrm{u}}}\left[\gamma' v w\right].$$

Here $\gamma'(a)$ denotes differentiation with respect to $s(a)$, and

$$\gamma'(a) = \left(\frac{h'_a}{h_a}\right)^2 \left[p + 1 - (2p+8)h'_a + 8\left(h'_a\right)^2\right],$$
$$\frac{\gamma'(a)^2}{h''(s(a))/h'_a} \leq 2(p+1)^2 M \lambda^{\mathrm{u}}.$$

For the latter inequality, the bracket has magnitude at most $p+1$ on $0 \leq h'_a \leq 1$, and $(h'_a/h_a)^2 = 4h'_a(1-h'_a)$, $h'_a \leq M\lambda^{\mathrm{u}}$, $h'_a/h_a \leq 1$. Centering contracts $L^2(\boldsymbol{x}^{\mathrm{u}})$. By (E.2) and (E.4),

$$\left\| \frac{\mathrm{d}}{\mathrm{d}q} \boldsymbol{J}^{\mathrm{u}}(\boldsymbol{s} + q\boldsymbol{v})\boldsymbol{w} \Big|_{q=0} \right\|_{(\boldsymbol{x}^{\mathrm{u}})^{-1}} \leq \left[2\gamma_{\max}\sqrt{\gamma_{\max}/\lambda^{\mathrm{u}}} + \sqrt{2M}(p+1)\right] \sqrt{\boldsymbol{v}^\top \boldsymbol{H}^{\mathrm{u}} \boldsymbol{v}}\, \|\boldsymbol{w}\|_\infty .$$

Using both bounds in (E.3) and scaling gives a coefficient at most $(4 + \sqrt{2/(p+1)})/\sqrt{M} < 5$. Thus

$$\left\| \frac{\mathrm{d}}{\mathrm{d}q} \boldsymbol{J}(\boldsymbol{\theta} + q\boldsymbol{v}) \Big|_{q=0} \right\|_{\infty \to \boldsymbol{x}^{-1}} \leq 5\sqrt{\boldsymbol{v}^\top \boldsymbol{H}(\boldsymbol{\theta}) \boldsymbol{v}}. \tag{E.7}$$

Fix $\boldsymbol{\theta}_0, \boldsymbol{v} \in \mathbb{R}^m$ with $\|\boldsymbol{v}\|_\infty \leq \epsilon$, and set $\boldsymbol{\theta}_s = \boldsymbol{\theta}_0 + s\boldsymbol{v}$ for $s \in [0,1]$. Subscripts s in the next calculations denote evaluation on this segment; for example, $\boldsymbol{y}_s = \boldsymbol{y}(\boldsymbol{\theta}_s)$ and $\boldsymbol{J}_s = \boldsymbol{J}(\boldsymbol{\theta}_s)$. Fenchel duality gives the two directed integrals

$$D_\psi(\boldsymbol{y}_1, \boldsymbol{y}_0) = \int_0^1 s \boldsymbol{v}^\top \boldsymbol{H}_s \boldsymbol{v}\, \mathrm{d}s,$$
$$D_\psi(\boldsymbol{y}_0, \boldsymbol{y}_1) = \int_0^1 (1-s) \boldsymbol{v}^\top \boldsymbol{H}_s \boldsymbol{v}\, \mathrm{d}s. \tag{E.8}$$

By (E.5) and (E.6), mass ratios are at most e^ϵ, coordinate ratios at most $e^{2\epsilon}$, and Hessian quadratic-form ratios at most $e^{9\epsilon/4}$. Integrating the two Hellinger bounds in (E.5) therefore yields

$$\|\sqrt{\boldsymbol{x}_1} - \sqrt{\boldsymbol{x}_0}\|_2 \leq \epsilon e^{\epsilon/2} \sqrt{\lambda_1},$$
$$\sqrt{\lambda_1}\, \|\sqrt{\boldsymbol{x}_1} - \sqrt{\boldsymbol{x}_0}\|_2^2 \leq \frac{e^{\epsilon/2}}{2} \int_0^1 \boldsymbol{v}^\top \boldsymbol{H}_s \boldsymbol{v}\, \mathrm{d}s \leq e^{11\epsilon/4} D_\psi(\boldsymbol{y}_1, \boldsymbol{y}_0).$$

Since $x_0(a)/x_1(a) \leq e^{2\epsilon}$,

$$\sqrt{\lambda_1}\, \|\boldsymbol{x}_1 - \boldsymbol{x}_0\|_{\boldsymbol{x}_1^{-1}}^2 \leq (1+e^\epsilon)^2 e^{11\epsilon/4} D_\psi(\boldsymbol{y}_1, \boldsymbol{y}_0) \leq 5 D_\psi(\boldsymbol{y}_1, \boldsymbol{y}_0)$$

when $\epsilon \leq 1/25$. For $\overline{\boldsymbol{J}} = \eta \int_0^1 \boldsymbol{J}_s\, \mathrm{d}s$ and every $\boldsymbol{w} \in \mathbb{R}^m$, Jensen's inequality in the endpoint Fisher norm gives

$$\begin{aligned} \left\|\overline{\boldsymbol{J}}\boldsymbol{w}\right\|_{\boldsymbol{x}_1^{-1}}^2 &\leq \eta^2 \int_0^1 e^{2\epsilon} \left\|\boldsymbol{J}_s \boldsymbol{w}\right\|_{\boldsymbol{x}_s^{-1}}^2 \mathrm{d}s \\ &\leq 4e^{5\epsilon} \eta^2 \lambda_1 \|\boldsymbol{w}\|_{\boldsymbol{x}_1}^2 \leq 5\eta^2 \lambda_1 \|\boldsymbol{w}\|_{\boldsymbol{x}_1}^2 . \end{aligned}$$

Integration by parts gives

$$\overline{\boldsymbol{J}} - \eta \boldsymbol{J}_1 = -\eta \int_0^1 s \frac{\mathrm{d}\boldsymbol{J}_s}{\mathrm{d}s}\, \mathrm{d}s,$$
$$\overline{\boldsymbol{J}} - \eta \boldsymbol{J}_0 = \eta \int_0^1 (1-s) \frac{\mathrm{d}\boldsymbol{J}_s}{\mathrm{d}s}\, \mathrm{d}s.$$

Weighted Cauchy–Schwarz, (E.7) and (E.8), and coordinate comparison give

$$\left\|\overline{\boldsymbol{J}} - \eta \boldsymbol{J}_1\right\|_{\infty \to \boldsymbol{x}_1^{-1}}^2 \leq \frac{25}{2} e^{2\epsilon} \eta^2 D_\psi(\boldsymbol{y}_1, \boldsymbol{y}_0),$$

$$\left\|\overline{\boldsymbol{J}} - \eta \boldsymbol{J}_0\right\|^2_{\infty\to \boldsymbol{x}_1^{-1}} \le \frac{25}{2} e^{17\epsilon/4} \eta^2 D_\psi(\boldsymbol{y}_1, \boldsymbol{y}_0).$$

Compare successive averages through their common endpoint and sum. The constant extension has zero movement, so no initial term occurs:

$$\sum_{t=1}^{T} \left\|\overline{\boldsymbol{J}}^{(t)} - \overline{\boldsymbol{J}}^{(t-1)}\right\|^2_{\infty\to(\boldsymbol{x}^{(t)})^{-1}} \le 25\left(e^{17\epsilon/4} + e^{4\epsilon}\right)\eta^2 \mathcal{P}_{i,T} \le 60\eta^2 \mathcal{P}_{i,T}.$$

These are the asserted estimates. The numerical comparisons follow, for example, from $e^u \le (1-u)^{-1}$ for $0 \le u < 1$. □

Lemma D.3. *Fix $i \in [N]$. For the update in* (5.1)*, suppose $2\eta \left\|\boldsymbol{e}_i^{(t)}\right\|_\infty \le 1/10$ for every $t \in [T]$. Then*

$$\mathcal{R}_{i,T} \le \frac{r_i}{\eta} + 3\eta \sum_{t=1}^{T} (\lambda_i^{(t)})^{3/2} \left\|(\mathsf{F}\boldsymbol{g}_i)^{(t)}\right\|^2_{\boldsymbol{x}_i^{(t)}} - \frac{\mathcal{P}_{i,T}}{2\eta}. \tag{D.7}$$

Proof. Fix $i \in [N]$ and $T \ge 1$, suppress the player index, write $\psi = \psi_i$, and let ψ^* denote its Fenchel conjugate. Set $\boldsymbol{e}^{(0)} = 0$. The played score is $\boldsymbol{\theta}^{(t)} = \eta\left(\sum_{s=1}^{t} \boldsymbol{g}^{(s)} - \boldsymbol{e}^{(t)}\right)$. Append the terminal response $\boldsymbol{y}^{(T+1)} = \nabla\psi^*\left(\eta \sum_{s=1}^{T} \boldsymbol{g}^{(s)}\right)$. For any feasible comparator $\widehat{\boldsymbol{y}}$, the regularized-leader telescope gives

$$\begin{aligned}\sum_{t=1}^{T} \left\langle \boldsymbol{g}^{(t)}, \widehat{\boldsymbol{y}} - \boldsymbol{y}^{(t)}\right\rangle \le& \frac{\psi(\widehat{\boldsymbol{y}}) - \psi(\boldsymbol{y}^{(0)})}{\eta} - \frac{1}{\eta}\sum_{t=1}^{T+1} D_\psi(\boldsymbol{y}^{(t)}, \boldsymbol{y}^{(t-1)}) \\ &+ \sum_{t=1}^{T} \left\langle \boldsymbol{e}^{(t)}, \boldsymbol{y}^{(t+1)} - \boldsymbol{y}^{(t)}\right\rangle.\end{aligned}$$

Indeed, apply the first-order convexity inequality to the score increments $\boldsymbol{g}^{(t)} - (\boldsymbol{e}^{(t)} - \boldsymbol{e}^{(t-1)})$, with the additional terminal increment $\boldsymbol{e}^{(T)}$, and sum their Bregman remainders. Reindexing the terms containing $\boldsymbol{e}$ gives the displayed expression.

For each $t \in [T]$, set $\widetilde{\boldsymbol{y}}^{(t)} = \nabla\psi^*(\boldsymbol{\theta}^{(t)} + 2\eta\boldsymbol{e}^{(t)})$ and recall that $2\eta\left\|\boldsymbol{e}^{(t)}\right\|_\infty \le 1/10$. For $s \in [0,1]$, set $\boldsymbol{\theta}_s = \boldsymbol{\theta}^{(t)} + 2s\eta\boldsymbol{e}^{(t)}$, $\lambda_s = \lambda(\boldsymbol{\theta}_s)$, and $\boldsymbol{x}_s = \boldsymbol{x}(\boldsymbol{\theta}_s)$. By (E.5), for every $s \in [0,1]$ and $a \in [m_i]$,

$$\lambda_s^{3/2} x_s(a) \le e^{7\eta\left\|\boldsymbol{e}^{(t)}\right\|_\infty} (\lambda^{(t)})^{3/2} x^{(t)}(a).$$

The Hessian bound in (E.5) and the directed integral in (E.8) therefore give

$$D_\psi(\boldsymbol{y}^{(t)}, \widetilde{\boldsymbol{y}}^{(t)}) \le 4e^{7\eta\left\|\boldsymbol{e}^{(t)}\right\|_\infty} \eta^2 (\lambda^{(t)})^{3/2} \left\|\boldsymbol{e}^{(t)}\right\|^2_{\boldsymbol{x}^{(t)}}.$$

The Bregman three-point identity yields

$$\begin{aligned}2\eta\left\langle \boldsymbol{e}^{(t)}, \boldsymbol{y}^{(t+1)} - \boldsymbol{y}^{(t)}\right\rangle =& D_\psi(\boldsymbol{y}^{(t+1)}, \boldsymbol{y}^{(t)}) + D_\psi(\boldsymbol{y}^{(t)}, \widetilde{\boldsymbol{y}}^{(t)}) - D_\psi(\boldsymbol{y}^{(t+1)}, \widetilde{\boldsymbol{y}}^{(t)}) \\ \le& D_\psi(\boldsymbol{y}^{(t+1)}, \boldsymbol{y}^{(t)}) + 4e^{7/20}\eta^2 (\lambda^{(t)})^{3/2} \left\|\boldsymbol{e}^{(t)}\right\|^2_{\boldsymbol{x}^{(t)}}.\end{aligned}$$

Since $\boldsymbol{y}^{(1)} = \boldsymbol{y}^{(0)}$, the remaining negative divergences sum to at most $-\mathcal{P}_{i,T}/(2\eta)$. The range is r_i, and $2e^{7/20} < 3$. Maximizing over the comparator proves the lemma. □

Lemma D.4. *Let H_t be coordinate Hilbert norms for $t \in [T]$, with the finite backward shift $\mathsf{L}_{\mathrm{out}}$ on $\ell^2(H)$. Let $R > \max\{1, \|\mathsf{L}_{\mathrm{out}}\|\}$ and let K have real Taylor coefficients and be analytic near $\{z : |z| \le R\}$. For real linear maps $\boldsymbol{J}^{(t)}$ from a fixed coordinate space into H_t, $t \in [T]$, set $\boldsymbol{J}^{(0)} = \boldsymbol{J}^{(1)}$. Write $\boldsymbol{J}$ for their pointwise action, set $\Delta\boldsymbol{J}^{(t)} := \boldsymbol{J}^{(t)} - \boldsymbol{J}^{(t-1)}$, and define $[K, \boldsymbol{J}]\boldsymbol{v} := K(\boldsymbol{J}\boldsymbol{v}) - \boldsymbol{J}(K\boldsymbol{v})$. For every real input sequence $\boldsymbol{v}$ extended by zero outside $[T]$, with $\|\boldsymbol{v}\|_\infty := \max_{t\in[T]} \left\|\boldsymbol{v}^{(t)}\right\|_\infty$,*

$$\|[K, \boldsymbol{J}]\boldsymbol{v}\|^2_{\ell^2(H)} \le \frac{\pi R^2 \max_{|z|=R} |K(z)|^2}{2(R^2 - \|\mathsf{L}_{\mathrm{out}}\|^2)(R^2 - 1)} \|\boldsymbol{v}\|^2_\infty \sum_{t=1}^{T} \left\|\Delta\boldsymbol{J}^{(t)}\right\|^2_{\infty\to H_t}. \tag{D.9}$$

Proof. Fix an integer $m \geq 1$ and a real linear map $\boldsymbol{J}_{\mathrm{cov}} : \mathbb{R}^m \to H$ into a real Hilbert space H. Write $\boldsymbol{J}_{\mathrm{cov},a}$ for its ath column, so $\boldsymbol{J}_{\mathrm{cov}}\boldsymbol{v} = \sum_{a=1}^{m} v(a)\boldsymbol{J}_{\mathrm{cov},a}$ for every $\boldsymbol{v} \in \mathbb{R}^m$. All coordinate sums in this proof run over $[m]$. For every real positive semidefinite matrix $C \in \mathbb{R}^{m\times m}$ with $C_{aa} \leq 1$ for $a \in [m]$,

$$\sum_{a,b} \langle \boldsymbol{J}_{\mathrm{cov},a}, \boldsymbol{J}_{\mathrm{cov},b} \rangle_H \, C_{ab} \leq \frac{\pi}{2} \, \|\boldsymbol{J}_{\mathrm{cov}}\|^2_{\infty\to H} \, . \tag{E.9}$$

Adding the diagonal matrix with entries $1 - C_{aa}$ increases the left-hand side, so assume $C_{aa} = 1$. Represent C as the Gram matrix of unit vectors $\boldsymbol{v}_a$, take a standard Gaussian vector Z in their span, and set $s_a = \operatorname{sign}\langle Z, \boldsymbol{v}_a\rangle$. Rotational symmetry in each two-dimensional span gives $\mathbb{E}\,[s_a s_b] = (2/\pi)\arcsin C_{ab}$. The entrywise matrix $\arcsin(C) - C$ is positive semidefinite: its power series has nonnegative coefficients, and each positive integer Schur power of C is a Gram matrix of tensor powers of the $\boldsymbol{v}_a$. Taking its trace against the Gram matrix of the $\boldsymbol{J}_{\mathrm{cov},a}$ gives

$$\frac{2}{\pi} \sum_{a,b} \langle \boldsymbol{J}_{\mathrm{cov},a}, \boldsymbol{J}_{\mathrm{cov},b} \rangle_H \, C_{ab} \leq \mathbb{E}\left[\|\boldsymbol{J}_{\mathrm{cov}}\boldsymbol{s}\|^2_H \right] \leq \|\boldsymbol{J}_{\mathrm{cov}}\|^2_{\infty\to H} \, .$$

The same inequality holds for a complex random vector with real covariance, because its squared norm under the complex extension of $\boldsymbol{J}_{\mathrm{cov}}$ has the same covariance expression.

Let L_{in} be the ordinary finite backward shift on input sequences, and $\mathsf{L}_{\mathrm{out}}$ the shift on the weighted output Hilbert space, both with zero at time 1. Both shifts are nilpotent. Cauchy's formula and the resolvent identity give

$$[K, \boldsymbol{J}] = \frac{1}{2\pi \mathrm{i}} \oint_{|z|=R} K(z)(zI - \mathsf{L}_{\mathrm{out}})^{-1}(\mathsf{L}_{\mathrm{out}}\boldsymbol{J} - \boldsymbol{J}\mathsf{L}_{\mathrm{in}})(zI - \mathsf{L}_{\mathrm{in}})^{-1} \, \mathrm{d}z.$$

The middle factor maps a sequence $\boldsymbol{v}$ to $-\Delta\boldsymbol{J}^{(t)}\boldsymbol{v}^{(t-1)}$. For an output test vector $\boldsymbol{w} \in \ell^2(H)$, orthogonality of the powers of $e^{\mathrm{i}\phi}$ gives

$$\begin{aligned}\frac{1}{2\pi} \int_{-\pi}^{\pi} \left\| \left((Re^{\mathrm{i}\phi} I - \mathsf{L}_{\mathrm{out}})^{-1} \right)^* \boldsymbol{w} \right\|^2 \mathrm{d}\phi &= \sum_{s=0}^{T-1} R^{-2s-2} \, \|(\mathsf{L}^*_{\mathrm{out}})^s \boldsymbol{w}\|^2 \\ &\leq \frac{\|\boldsymbol{w}\|^2}{R^2 - \|\mathsf{L}_{\mathrm{out}}\|^2}.\end{aligned}$$

For $t \in [T]$ and $\phi \in [-\pi, \pi]$, define $\boldsymbol{z}_t(\phi) = ((Re^{\mathrm{i}\phi} I - \mathsf{L}_{\mathrm{in}})^{-1}\boldsymbol{v})^{(t-1)}$. This vector has real covariance

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} z_t(\phi)(a)\overline{z_t(\phi)(b)} \, \mathrm{d}\phi = \sum_{s\geq 0} R^{-2s-2} v^{(t-1-s)}(a) v^{(t-1-s)}(b).$$

Its diagonal entries are at most $\|\boldsymbol{v}\|^2_\infty / (R^2 - 1)$; zero extension makes each sum finite. Apply (E.9) with $\boldsymbol{J}_{\mathrm{cov}} = \Delta\boldsymbol{J}^{(t)}$ in H_t, and then sum over time:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left\| (\mathsf{L}_{\mathrm{out}}\boldsymbol{J} - \boldsymbol{J}\mathsf{L}_{\mathrm{in}})(Re^{\mathrm{i}\phi} I - \mathsf{L}_{\mathrm{in}})^{-1}\boldsymbol{v} \right\|^2_{\ell^2(H)} \, \mathrm{d}\phi \leq \frac{\pi \, \|\boldsymbol{v}\|^2_\infty}{2(R^2 - 1)} \sum_t \left\| \Delta\boldsymbol{J}^{(t)} \right\|^2_{\infty\to H_t} .$$

Pair the contour integral with $\boldsymbol{w}$, use these two estimates and Cauchy–Schwarz in ϕ, and take the supremum over unit test vectors. The resulting coefficient is

$$\frac{\pi R^2 \max_{|z|=R} |K(z)|^2}{2(R^2 - \|\mathsf{L}_{\mathrm{out}}\|^2)(R^2 - 1)},$$

as claimed. The argument uses analyticity on the disk and finite shifts, so it applies to nonrational filters and every finite horizon. □

Lemma D.5. *For every $N \geq 2$, set $\rho = 1 - 1/N$ and $\delta(z) = (1 - z)/(1 - \rho z)$. Then*

$$\left\| 1 - (1 - z)\delta(z)^{N-1} \right\|_{\ell^1} \leq 48.$$

Proof. Write $\rho = 1 - 1/N$ and $\delta(z) = (1-z)/(1-\rho z)$, with $N \geq 2$. On the unit circle the maximum of $|\delta(z)|$ occurs at $z = -1$, and hence

$$\max_{|z|=1} |\delta(z)|^{N-1} = \left(1 - \frac{1}{2N}\right)^{-(N-1)} \leq e^{2/3} < 2.$$

For $t \in \mathbb{R}$, set $F_\circ(t) = (1 - e^{\mathrm{i}t})\delta(e^{\mathrm{i}t})^{N-1}$. Differentiation gives

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}z}\left[(1-z)\delta(z)^{N-1}\right] =& -\delta(z)^{N-1}\left[1 + \frac{N-1}{N}(1-\rho z)^{-1}\right], \\
|F_\circ(t)| \leq& 2\,\mathrm{dist}(t, 2\pi\mathbb{Z}), \\
|F_\circ'(t)| \leq& 2 + \frac{\sqrt{2}\pi}{\mathrm{dist}(t, 2\pi\mathbb{Z})} \quad (t \notin 2\pi\mathbb{Z}).
\end{aligned}$$

The last bound uses $\left|1 - \rho e^{\mathrm{i}t}\right| \geq 2\sqrt{\rho}\ \mathrm{dist}(t, 2\pi\mathbb{Z})/\pi$. Fix $0 < u \leq 1$. On $\mathrm{dist}(t, 2\pi\mathbb{Z}) \leq 2\sqrt{u}$, the 1-Lipschitz property of distance gives $|F_\circ(t+u) - F_\circ(t)| \leq 2(2\,\mathrm{dist}(t, 2\pi\mathbb{Z}) + u)$. Therefore its contribution to the squared normalized L^2 norm is at most

$$\frac{4}{\pi}\int_0^{2\sqrt{u}} (2t+u)^2\,\mathrm{d}t \leq \frac{248}{3\pi}u^{3/2}.$$

On the complement, the distance along the intervening arc is at least $\mathrm{dist}(t, 2\pi\mathbb{Z}) - u \geq \mathrm{dist}(t, 2\pi\mathbb{Z})/2$. Its contribution is at most

$$\frac{u^2}{\pi}\int_{2\sqrt{u}}^{\pi} \left(2 + \frac{2\sqrt{2}\pi}{t}\right)^2 \mathrm{d}t \leq \left(4 + 4\pi + \frac{4\sqrt{2}\pi}{e}\right) u^{3/2},$$

where $\max_{0<s\leq 1} s\log(\pi/(2s)) = \pi/(2e)$ bounds the logarithmic term. Thus

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |F_\circ(t+u) - F_\circ(t)|^2\ \mathrm{d}t \leq 50u^{3/2}.$$

Let f_s be the Taylor coefficients of $(1-z)\delta(z)^{N-1}$. For an integer $L \geq 1$, use $u = 1/L$. On $L \leq s < 2L$, $\left|e^{\mathrm{i}s/L} - 1\right|^2 \geq 4\sin^2(1/2) > 9/10$. Parseval's identity and Cauchy–Schwarz then give

$$\begin{aligned}
\sum_{s=L}^{2L-1} |f_s|^2 \leq& \frac{500}{9}L^{-3/2}, \\
\sum_{s\geq 1} |f_s| \leq& \sum_{j\geq 0} \sqrt{500/9}\, 2^{-j/4} = \frac{\sqrt{500/9}}{1 - 2^{-1/4}} < 48.
\end{aligned}$$

The constant coefficient is $f_0 = 1$. It cancels in $I - \mathsf{D}\delta^{N-1}$, proving the lemma. □

F ADVERSARIAL REGRET

For every player $i \in [N]$, define the threshold from Theorem 4.2 by

$$\overline{\mathrm{Reg}_i} \coloneqq M_i\left(2000Nr_i + \sum_{j=1}^{N} r_j + \frac{3}{40N}\right).$$

Set $\mathrm{Reg}_i(0) = 0$ and run Algorithm 1 with the learning rate at timestep $t \geq 1$ chosen as

$$\eta_i^{(t)} \coloneqq \begin{cases} 1/(2000N), & \mathrm{Reg}_i(t-1) \leq \overline{\mathrm{Reg}_i}, \\ \dfrac{1}{10}\sqrt{r_i/t}, & \mathrm{Reg}_i(t-1) > \overline{\mathrm{Reg}_i}, \end{cases} \tag{F.1}$$

Thus the response in (4.7) uses $\eta_i^{(t)}$ in place of η, with the same regularizer, predictor, and cumulative feedback. The rate uses only observed feedback and returns to $1/(2000N)$ whenever the regret is at or below the threshold. This rule is deterministic, uncoupled, and independent of the horizon.

Proposition F.1. *If every player uses* (F.1) *in a fixed game, the trajectory coincides with that of Algorithm 1 with $\eta = 1/(2000N)$, and Theorem 4.2 holds unchanged. For every player $i \in [N]$ and every sequence of payoff vectors $\boldsymbol{u}_i^{(t)} \in [0,1]^{m_i}$, possibly chosen adaptively, the same update satisfies, simultaneously for every $T \geq 1$,*

$$[\mathrm{Reg}_i(T)]_+ \leq M_i \left(\overline{\mathrm{Reg}}_i + 1 + 30\sqrt{r_i T} \right) \leq O\left(N \log m_{\max} + \sqrt{T \log m_i} \right),$$

where the implicit constant is universal.

Proof. In self-play, all players use the constant rate at timestep 1 because $\mathrm{Reg}_i(0) = 0$. For every $t \geq 2$, if play coincides through timestep $t-1$, then Theorem 4.2 gives $\mathrm{Reg}_i(t-1) \leq \overline{\mathrm{Reg}}_i$ for every $i \in [N]$, so the rates and strategies also coincide at timestep t. Induction proves the self-play claim.

Fix $i \in [N]$ and an arbitrary payoff sequence. For every $\eta > 0$ and $\boldsymbol{z} \in \mathbb{R}^{m_i}$, define

$$\psi_{i,\eta}^*(\boldsymbol{z}) \coloneqq \max_{\boldsymbol{y} \in \mathrm{dom}\, \psi_i} \left\{ \langle \boldsymbol{z}, \boldsymbol{y} \rangle - \frac{\psi_i(\boldsymbol{y})}{\eta} \right\}.$$

Its gradient is $\boldsymbol{y}_i(\eta \boldsymbol{z})$. By (E.5) and lemma 4.1, the positive-semidefinite Hessian of the Fenchel conjugate ψ_i^* has operator norm at most 2 from ℓ_∞ to ℓ_1 at every score. Consequently, $\nabla \psi_{i,\eta}^*$ is 2η-Lipschitz in these norms. Since $-r_i \leq \psi_i \leq 0$ and every feasible $\boldsymbol{y}$ is nonnegative with $\|\boldsymbol{y}\|_1 \leq 1$, for every $\boldsymbol{z} \in \mathbb{R}^{m_i}$ and $0 < \eta' \leq \eta$,

$$\frac{[\max_{a_i \in [m_i]} z(a_i)]_+}{M_i} \leq \psi_{i,\eta}^*(\boldsymbol{z}) \leq [\max_{a_i \in [m_i]} z(a_i)]_+ + \frac{r_i}{\eta}, \tag{F.2}$$

$$\psi_{i,\eta'}^*(\boldsymbol{z}) \leq \psi_{i,\eta}^*(\boldsymbol{z}) + r_i \left(\frac{1}{\eta'} - \frac{1}{\eta} \right). \tag{F.3}$$

The lower bound uses the feasible comparators $\mathbf{0}$ and the pure strategies scaled by $1/M_i$.

By Lemma D.5, $\left\| \widehat{\boldsymbol{g}}_i^{(t)} \right\|_\infty \leq 48$ for every $t \geq 1$. Since $\left\langle \boldsymbol{y}_i^{(t)}, \boldsymbol{g}_i^{(t)} \right\rangle = 0$ and $\left\| \boldsymbol{g}_i^{(t)} \right\|_\infty \leq 1$, smoothness and (4.7) yield

$$\begin{aligned} &\psi_{i,\eta_i^{(t)}}^* \left(\sum_{s=1}^{t} \boldsymbol{g}_i^{(s)} \right) - \psi_{i,\eta_i^{(t)}}^* \left(\sum_{s=1}^{t-1} \boldsymbol{g}_i^{(s)} \right) \\ &\leq \left\langle \nabla \psi_{i,\eta_i^{(t)}}^* \left(\sum_{s=1}^{t-1} \boldsymbol{g}_i^{(s)} \right) - \boldsymbol{y}_i^{(t)}, \boldsymbol{g}_i^{(t)} \right\rangle + \eta_i^{(t)} \\ &\leq \left(1 + 2 \left\| \widehat{\boldsymbol{g}}_i^{(t)} \right\|_\infty \right) \eta_i^{(t)} \leq 97 \eta_i^{(t)}. \end{aligned} \tag{F.4}$$

Fix $T \geq 1$ with $\mathrm{Reg}_i(T) > \overline{\mathrm{Reg}}_i$, since otherwise the asserted bound already holds. Let $t_i \coloneqq \max \left\{ t \in \{0, \ldots, T-1\} : \mathrm{Reg}_i(t) \leq \overline{\mathrm{Reg}}_i \right\}$ be the last timestep at or below the threshold. The bound on each payoff gives $\mathrm{Reg}_i(t_i + 1) \leq \overline{\mathrm{Reg}}_i + 1$, which proves the claim if $T = t_i + 1$. Otherwise, (F.1) gives $\eta_i^{(t)} = \sqrt{r_i/t}/10$ for every integer $t_i + 2 \leq t \leq T$. These rates are nonincreasing. Summing (F.3) and (F.4) over this interval and using (F.2) at its endpoints gives

$$\begin{aligned} \frac{\mathrm{Reg}_i(T)}{M_i} &\leq \psi_{i,\eta_i^{(T)}}^* \left(\sum_{s=1}^{T} \boldsymbol{g}_i^{(s)} \right) \\ &\leq \psi_{i,\eta_i^{(t_i+2)}}^* \left(\sum_{s=1}^{t_i+1} \boldsymbol{g}_i^{(s)} \right) + r_i \left(\frac{1}{\eta_i^{(T)}} - \frac{1}{\eta_i^{(t_i+2)}} \right) + 97 \sum_{t=t_i+2}^{T} \eta_i^{(t)} \\ &\leq \overline{\mathrm{Reg}}_i + 1 + \frac{r_i}{\eta_i^{(T)}} + 97 \sum_{t=t_i+2}^{T} \eta_i^{(t)} \end{aligned}$$

$$\leq \overline{\mathrm{Reg}_i} + 1 + \left(10 + \frac{97}{5}\right)\sqrt{r_i T}.$$

The last inequality uses $\sum_{t=1}^{T} t^{-1/2} \leq 2\sqrt{T}$. Since $M_i \leq e$ and $r_i \leq O(\log m_i)$, the stated bound follows. □