Cautious Optimism: A Meta-Algorithm for Near-Constant Regret in General Games

Ashkan Soleymani ASHKANSO@MIT.EDU
MIT

Georgios Piliouras GPIL@GOOGLE.COM
Google DeepMind

Gabriele Farina GFARINA@MIT.EDU
MIT

Abstract

Recent work [Soleymani et al., 2025] introduced a variant of Optimistic Multiplicative Weights Updates (OMWU) that adaptively controls the learning pace in a dynamic, non-monotone manner, achieving new state-of-the-art regret minimization guarantees in general games. In this work, we demonstrate that no-regret learning acceleration through adaptive pacing of the learners is not an isolated phenomenon. We introduce *Cautious Optimism*, a framework for substantially faster regularized learning in general games. Cautious Optimism takes as input any instance of Follow-the-Regularized-Leader (FTRL) and outputs an accelerated no-regret learning algorithm by pacing the underlying FTRL with minimal computational overhead. Importantly, we retain uncoupledness (learners do not need to know other players' utilities). Cautious Optimistic FTRL achieves near-optimal $O_T(\log T)$ regret in diverse self-play (mixing-and-matching regularizers) while preserving the optimal $O(\sqrt{T})$ regret in adversarial scenarios. In contrast to prior works (e.g. Syrgkanis et al. [2015], Daskalakis et al. [2021]), our analysis does not rely on monotonic step-sizes, showcasing a novel route for fast learning in general games.

The full version of this paper is available at arXiv:2506.05005.

1. Introduction

No-regret learning lies at the core of the interplay between online learning and game theory. While classical no-regret algorithms guarantee $O_T(\sqrt{T})$ regret in adversarial settings [Cesa-Bianchi and Lugosi, 2006], this worst-case bound is overly pessimistic in games, where agents repeatedly interact in structured, predictable environments. A central open question is therefore: *what are the fastest no-regret algorithms in games?*

The optimism framework of Syrgkanis et al. [2015] took the first major step, showing that optimistic regularized learning achieves $O_T(T^{1/4})$ regret in self-play. This result sparked a series of work [Chen and Peng, 2020, Daskalakis et al., 2021, Farina et al., 2022, Soleymani et al., 2025] trying to achieve faster regret rates in games by designing specialized, tailor-made algorithms. Despite the significant exponential improvement to $O_T(\log T)$ in the most recent works [Daskalakis et al., 2021, Farina et al., 2022, Soleymani et al., 2025], this progress remains confined to single-instance algorithms and lacks the generality of the *Optimism* framework of Syrgkanis et al. [2015]. In this work, we aim to bridge this gap.

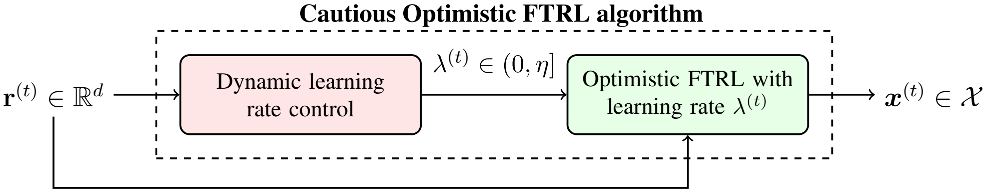
Cautious Optimistic FTRL algorithm
$\mathbf{r}^{(t)} \in \mathbb{R}^d$
Dynamic learning rate control
$\lambda^{(t)} \in (0, \eta]$
Optimistic FTRL with learning rate $\lambda^{(t)}$
$\boldsymbol{x}^{(t)} \in \mathcal{X}$

Figure 1: Dynamics of Cautious Optimistic Follow-the-Regularized-Leader (COFTRL) Algorithms. COFTRL takes as input an instance of an OFTRL algorithm and equips it with a dynamic learning rate control mechanism that nonmonotonically adjusts the learning rate of the underlying OFTRL instance. We prove that this simple and lightweight overhead on top of OFTRL leads to exponentially faster convergence guarantees for no-regret learning in games [Syrgkanis et al., 2015], for a broad class of regularizers.

Our contributions. We introduce *Cautious Optimism*, a meta-algorithm that takes as input an instance of the Follow-the-Regularized-Leader (FTRL) algorithm (typically achieving $O_T(T^{1/4})$ regret[1]) and produces a substantially accelerated no-regret learning algorithm, Cautious Optimistic Follow-the-Regularized-Leader (COFTRL), with $O_T(\log T)$ regret. This acceleration is achieved by dynamically adjusting the pace of the underlying FTRL method, using a structured, nonmonotonic adaptation of the learning rate that enables significantly faster convergence while maintaining minimal computational overhead. In Figure 1, we summarize the structure of COFTRL.

Our contributions can be summarized as follows:

- **Exponential acceleration.** Our work provides the first comprehensive characterization of exponentially faster regret minimization for regularized learning in general games. COFTRL achieves near-optimal $O_T(\log T)$ regret in games, while simultaneously maintaining the optimal $O_T(\sqrt{T})$ bound in adversarial environments.

- **Generality and unification.** Our framework retains the generality of Syrgkanis et al. [2015], and recovers LRL-OFTRL [Farina et al., 2022] and COMWU [Soleymani et al., 2025] as special cases, providing a unified explanation for their fast convergence. Unlike their tailored analyses, our guarantees follow directly from a general unified principle of adaptive learning rate control.

- **New state-of-the-art instances.** We construct two additional instances COFTRL algorithms that match the best known regret guarantees $O(n \log^2 d \log T)$ [Soleymani et al., 2025]. Please see Table 1. Moreover, COFTRL supports *diverse self-play*, where players mix and match their choice of regularizers for FTRL, without requiring all players to use the exact same instance of COFTRL.

- **Best-of-both-worlds guarantees.** COFTRL attains $O_T(\log T)$ individual regret and $O_T(1)$ social regret simultaneously, unlike prior optimistic methods which require different learning rate choices for these goals [Syrgkanis et al., 2015, Corollary 8 and 12].

[1] For general OFTRL algorithms [Syrgkanis et al., 2015].

Techniques. At the technical level, we introduce a relaxed notion of Lipschitz continuity, termed *intrinsic Lipschitzness*, which quantifies changes in the regularizer value through its *Bregman divergence*. COFTRL applies to any FTRL algorithm with an *intrinsically Lipschitz* regularizer. As we will demonstrate, intrinsic Lipschitzness is a mild condition, satisfied by a broad class of regularizers, including all Lipschitz continuous regularizers. We show that the dynamics of COFTRL are equivalent to an instance of OFTRL on the lifted space $(0,1]\Delta^d$ with a specific composite regularizer obtained as the sum of a strongly convex part and a nonconvex transformation of the original regularizer used by the underlying FTRL. We prove strong convexity of the composite regularizer in a specific regime of hyperparameters and provide technical steps to convert the regret analysis of the resulting OFTRL on the lifted space $(0,1]\Delta^d$ to that of COFTRL. As a byproduct of the strong convexity of the regularizer along the rays, we prove multiplicative stability of the dynamic learning rates to ensure smooth evolution of the no-regret learning dynamics. These techniques hold broadly for any intrinsic Lipschitz regularizer and are not constrained by the structural properties of the learning rate control problem.

2. Cautious Optimism

Cautious Optimism is a variant of the Optimistic Follow-the-Regularized-Leader (OFTRL) algorithms, but with non-monotone, adaptive adjustment of the learning rate based on the regret accumulated up to the current iteration t. As depicted in Figure 1, our framework can be seen as a module that takes an instance of OFTRL as input and accelerates its convergence guarantees substantially by adding a learning rate control problem on top. In the standard version of OFTRL with regularizer ψ, the actions of the play are picked according to

$$\boldsymbol{x}^{(t)} \leftarrow \operatorname*{arg\,max}_{\boldsymbol{x}\in\mathcal{X}} \left\{ \lambda^{(t)} \langle \mathbf{r}^{(t)}, \boldsymbol{x} \rangle - \psi(\boldsymbol{x}) \right\}, \tag{1}$$

where learning rate at time t is represented by $\lambda^{(t)} > 0$, and $\mathbf{r}^{(t)}$ denotes the vector containing the accumulated optimistically-corrected regrets for each action up to time t. This is given by,

$$\mathbf{r}^{(t)}[k] \coloneqq (\boldsymbol{\nu}^{(t-1)}[k] - \langle \boldsymbol{\nu}^{(t-1)}, \boldsymbol{x}^{(t-1)} \rangle) + \sum_{\tau=1}^{t-1} \left[\boldsymbol{\nu}^{(\tau)}[k] - \langle \boldsymbol{\nu}^{(\tau)}, \boldsymbol{x}^{(\tau)} \rangle \right],$$

for all $k \in \mathcal{A}$. To simplify further, let us define the corrected reward signal as $\boldsymbol{u}^{(t)} \coloneqq \boldsymbol{\nu}^{(t)} - \langle \boldsymbol{\nu}^{(t)}, \boldsymbol{x}^{(t)} \rangle \mathbf{1}_d$, while the accumulated signal is expressed as $\boldsymbol{U}^{(t)} \coloneqq \sum_{\tau=1}^{t-1} \left[\boldsymbol{\nu}^{(\tau)} - \langle \boldsymbol{\nu}^{(\tau)}, \boldsymbol{x}^{(\tau)} \rangle \mathbf{1}_d \right]$.

The celebrated work of Syrgkanis et al. [2015] showed that when all players in a game employ OFTRL with a fixed learning rate $\lambda^{(t)} = \eta$, the maximum regret accumulated by the players grows at most as $O_T(T^{1/4})$.[2] In this paper, we improve this result exponentially to an $O_T(\log T)$ dependence for a general class of regularizers, building on the *dynamic learning rate control* idea of Soleymani et al. [2025] for OMWU. Unlike conventional methods in optimization and learning that enforce a monotonically decreasing learning rate schedule, our approach allows for more flexible, non-monotone adjustments based on the learner's performance. In particular, we introduce a *universal dynamic learning control* mechanism that *deliberately slows down the learning* process of the underlying OFTRL when *regret becomes excessively negative*.

[2] Only the dependence on the time horizon T is shown. For details, please refer to Table 1.

Method	**Regret in Games**	**Adversarial Regret**	**General Learners**
OFTRL / OOMD [Syrgkanis et al., 2015]	$O(\sqrt{n}\,\mathfrak{R}(d)T^{1/4})$	$\widetilde{O}(\sqrt{T\log d})$	✓
COFTRL **[This paper]**	$O(n\Gamma(d)\log T)$	$\widetilde{O}(\sqrt{T\log d})$	✓
OMWU [Chen and Peng, 2020]	$O(n\log^{5/6} d\,T^{1/6})$ †	$\widetilde{O}(\sqrt{T\log d})$	✗
OMWU [Daskalakis et al., 2021]	$O(n\log d\log^4 T)$	$\widetilde{O}(\sqrt{T\log d})$	✗
Clairvoyant MWU [Piliouras et al., 2022]	$O(n\log d)$ for a subsequence only ‡	No guarantees	✗
LRL-OFTRL [Farina et al., 2022] **[≡ `COFTRL` w/ log regularizer]**	$O(n\,d\log T)$	$\widetilde{O}(\sqrt{T\log d})$	✗
COMWU [Soleymani et al., 2025] **[≡ `COFTRL` w/ negative entropy]**	$O(n\log^2 d\log T)$	$\widetilde{O}(\sqrt{T\log d})$	✗
`COFTRL` with ℓ_{p^*} **[This paper]**	$O(n\log^2 d\log T)$	$\widetilde{O}(\sqrt{T\log d})$	✗
`COFTRL` with q^*-Tsallis entropy **[This paper]**	$O(n\log^2 d\log T)$	$\widetilde{O}(\sqrt{T\log d})$	✗

Table 1: Comparison of existing no-regret learning algorithms in general games. We define n as the number of players, T as the number of game repetitions, and d as the number of available actions. For simplicity, dependencies on smoothness and utility range are omitted. † Applicable only to two-player games ($n = 2$). ‡ Unlike other algorithms, Clairvoyant MWU (CMWU) does not guarantee sublinear regret for its full sequence of iterates. Instead, after T iterations, only a subsequence of length $\Theta(T/\log T)$ achieves the regret bound stated in the table.

In self-play settings, the predictability of the players' actions due to the smooth evolution of the learning dynamics guarantees faster convergence compared to adversarial settings. This observation motivates us to consider the convergence of the learning dynamics of multiple players as a whole, rather than at an individual level. Hence, a learner performing exceedingly well can create an imbalance that hinders others from keeping pace. Consequently, aiming for harmonic learning among players during self-play seems a natural approach to better exploit the predictability of the dynamics. While hindering learning when a player is performing exceptionally well may seem counterintuitive and even unfavorable at first glance, it helps maintain a balance among the players, thereby improving the performance of the hindered player in the long run.

There is another way to justify at the conceptual level why such ideas lead to faster convergence. We know that no-regret dynamics converge to Coarse Correlated Equilibria at a rate dictated by the

Algorithm 1: Cautious Optimistic FTRL (COFTRL)

```
Data: Learning rate $\eta$, parameters $\alpha$
Set $\boldsymbol{U}^{(1)}, \boldsymbol{u}^{(0)} \leftarrow \mathbf{0} \in \mathbb{R}^d$
for $t = 1, 2, \ldots, T$ do
    /* Optimism */
    Set $\mathbf{r}^{(t)} \leftarrow \boldsymbol{U}^{(t)} + \boldsymbol{u}^{(t-1)}$
    /* Dynamic Learning Rate Control */
    Set $\lambda^{(t)} \leftarrow \arg\max_{\lambda \in (0,\eta]} \left\{ \alpha \log \lambda + \psi^*_{\mathcal{X}}(\lambda \mathbf{r}^{(t)}) \right\}$
    /* OFTRL with Dynamic Learning Rate */
    Set $\boldsymbol{x}^{(t)} \leftarrow \arg\max_{\boldsymbol{x} \in \Delta^d} \left\{ \lambda^{(t)} \langle \mathbf{r}^{(t)}, \boldsymbol{x} \rangle - \psi(\boldsymbol{x}) \right\}$
    Play strategy $\boldsymbol{x}^{(t)}$
    Observe $\boldsymbol{\nu}^{(t)} \in \mathbb{R}^d$
    /* Empirical Cumulated Regrets */
    Set $\boldsymbol{u}^{(t)} \leftarrow \boldsymbol{\nu}^{(t)} - \langle \boldsymbol{\nu}^{(t)}, \boldsymbol{x}^{(t)} \rangle \mathbf{1}_d$
    Set $\boldsymbol{U}^{(t+1)} \leftarrow \boldsymbol{U}^{(t)} + \boldsymbol{u}^{(t)}$
end
```

worst-performing player—the one with the highest regret. Thus, ensuring a balanced performance among players naturally accelerates convergence [Soleymani et al., 2025].

Intrinsic Lipschitzness. Cautious Optimism works for a broad class of convex regularizers, which we term *intrinsically Lipschitz.*

Definition 2.1 *Let $\psi : \Delta^d \to \mathbb{R}$ be an arbitrary convex regularizer for the simplex, we call ψ, γ-intrinsically Lipschitz (γ-IL) if for all $\boldsymbol{x}', \boldsymbol{x} \in \Delta^d$,*

$$\left|\psi(\boldsymbol{x}') - \psi(\boldsymbol{x})\right|^2 \leq \gamma D_\psi(\boldsymbol{x}' \parallel \boldsymbol{x}).$$

This condition is broadly applicable and satisfied by many regularizers. In particular, any Lipschitz and strongly convex function ψ is intrinsically Lipschitz, making intrinsic Lipschitzness strictly weaker than standard Lipschitzness in this context. As illustrated in Table 2, several widely used regularizers—including negative entropy and negative Tsallis entropy—are intrinsically Lipschitz but not Lipschitz in the standard sense.

Regret Bounds. Cautious achieves a individual regret of $O(n\Gamma_\psi(d)\log T)$ and social regret of $O_T(1)$ simultaneously in self-play and the optimal $O(\sqrt{T \log d})$ rate in adversarial settings, where $\Gamma_\psi(d)$ depends on the properties of the chosen regularizer ψ. We present the structure of Cautious Optimistic FTRL (COFTRL) in Figure 1 and its pseudocode in Algorithm 1. Please see Table 2 for different instantiations of COFTRL and the resulting regret bounds. We summarize our results as follows.

Theorem 2.2 *If all players $i \in [n]$ follow* `COFTRL` *with a γ-IL and μ-strongly convex*[3] *regularizer ψ, and a small enough learning rate cap $\eta = O_T(1)$, then the individual regret for each player $i \in [n]$ and social regret, we have,*

$$\mathrm{Reg}_i^{(T)} = O(n\Gamma_\psi(d)\log T), \textit{ and } \sum_{i\in[n]} \mathrm{Reg}_i^{(T)} = O_T(1),$$

where $\Gamma_\psi(d) = \gamma/\mu$ and the algorithm for each player $i \in [n]$ is adaptive to adversarial utilities, i.e., the regret that each player incurs is $\mathrm{Reg}_i^{(T)} = O(\sqrt{T \log d})$.

We conclude this section with the direct corollary that if n players follow the uncoupled learning dynamics of `COFTRL` for T rounds in a general-sum multiplayer game with a finite set of d deterministic strategies per player, the resulting empirical distribution of play constitutes an $O\left(\frac{n\Gamma(d)\log T}{T}\right)$-approximate coarse correlated equilibrium (CCE) of the game.

3. Conclusion

Cautious Optimism builds on `FTRL` by introducing *non-monotone adaptive learning rate control*, effectively pacing the learners in a dynamic, state-dependent manner. This marks a fundamental shift from the traditional paradigm of online learning, optimization, and game theory, where step sizes are constant or monotonically decreasing. The move to adaptive, state-dependent step sizes is not merely cosmetic, it has major implications. Classical dynamical systems tools (e.g., Poincaré recurrence Barreira [2006], center-stable manifold theorem Perko [2013], period-three-implies-chaos Li and Yorke [1975]) apply primarily to autonomous smooth maps or flows. While such methods have been used in game-theoretic learning, this has either relied on idealized continuous-time systems Kleinberg et al. [2009], Mertikopoulos et al. [2018], Piliouras and Shamma [2014], or fixed-step dynamics that sacrifice strong regret guarantees Bailey and Piliouras [2018], Bailey et al. [2020], Chotibut et al. [2020], Katona et al. [2024], Piliouras and Yu [2023], Wibisono et al. [2022]. Cautious Optimism paves a way forward where we do not have to make such concessions. This framework shows that adaptive, non-monotone step sizes can retain black-box regret guarantees while opening new directions for analysis. This raises broad research questions: how such dynamics evolve, their relation to chaotic behavior, their continuous-time counterparts, extensions beyond external regret (e.g., swap regret), and connections to social welfare and strict equilibria.

References

James P. Bailey and Georgios Piliouras. Multiplicative weights update in zero-sum games. In *ACM Conference on Economics and Computation*, 2018.

James P Bailey, Gauthier Gidel, and Georgios Piliouras. Finite regret and cycles with fixed step-size via alternating gradient descent-ascent. In *Conference on Learning Theory*, pages 391–407. PMLR, 2020.

Luis Barreira. Poincare recurrence: old and new. In *XIVth International Congress on Mathematical Physics. World Scientific.*, pages 415–422, 2006.

[3] w.r.t. ℓ_1 norm.

Regularizer	**Formulation ψ**	**Strong convexity μ**	**(L)IL parameter γ**	**Regret**	**Globally IL**
Negative entropy	$\sum_{k=1}^{d} \boldsymbol{x}[k] \log \boldsymbol{x}[k]$	$\Omega(1)$	$O(\log^2 d)$	$O(n \log^2 d \log T)$	✓
Log	$-\sum_{k=1}^{d} \log \boldsymbol{x}[k]$	$\Omega(1)$	$O(d)$	$O(nd \log T)$	✗
Squared ℓ_2 norm	$\frac{1}{2}\|\boldsymbol{x}\|_2^2$	$\Omega(d^{-1})$	$O(1)$	$O(nd \log T)$	✓
Squared ℓ_p norm ($p \in (1, 2]$)	$\frac{1}{2}\|\boldsymbol{x}\|_p^2$	$\Omega\big((p-1)d^{2/p-2}\big)$	$O\Big(\frac{1}{p-1}\Big)$	$O\Big(n \frac{d^{2-2/p}}{(p-1)^2} \log T\Big)$	✓
Squared ℓ_{p^*} norm ($p^* = 1 + 1/\log d$)	$\frac{1}{2}\|\boldsymbol{x}\|_{p^*}^2$	$\Omega\big((p^*-1)d^{2/p^*-2}\big)$	$O\Big(\frac{1}{p^*-1}\Big)$	$O(n \log^2 d \log T)$	✓
$1/2$-Tsallis entropy	$2\Big(1 - \sum_{k=1}^{d} \sqrt{\boldsymbol{x}[k]}\Big)$	$\Omega(1)$	$O(\sqrt{d})$	$O(n\sqrt{d} \log T)$	✓
q-Tsallis entropy ($q \in (0, 1)$)	$\frac{1}{1-q}\Big(1 - \sum_{k=1}^{d} \sqrt{\boldsymbol{x}[k]}\Big)$	$\Omega(q)$	$O\Big(\frac{d^{1-q}}{(1-q)^2}\Big)$	$O\Big(n \frac{d^{1-q}}{q(1-q)^2} \log T\Big)$	✓
q^*-Tsallis entropy ($q^* = 1 - 1/\log d$)	$\frac{1}{1-q^*}\Big(1 - \sum_{k=1}^{d} \sqrt{\boldsymbol{x}[k]}\Big)$	$\Omega(q^*)$	$O\Big(\frac{d^{1-q^*}}{(1-q^*)^2}\Big)$	$O\big(n \log^2 d \log T\big)$	✓
L-Lipschitz and μ-strongly convex	general	μ	$O\Big(\frac{L^2}{\mu}\Big)$	$O\Big(n \frac{L^2}{\mu^2} \log T\Big)$	✓

Table 2: Various examples of appropriate regularizers ψ for Cautious Optimism, which are (locally) intrinsically Lipschitz and strongly convex, along with the corresponding regret rates of `COFTRL` with each choice of ψ. `COFTRL` instantiated with negative entropy, the squared ℓ_{p^*} norm (for an appropriate choice of $p^* = 1 + 1/\log d$), or the q^*-Tsallis entropy (for an appropriate choice of $q^* = 1 - 1/\log d$) leads to new state-of-the-art no-regret algorithms in games. The parameter μ represents the strong convexity parameter with respect to the ℓ_1 norm. The log regularizer is $O(d)$-LIL, whereas the other examples in this table are intrinsically Lipschitz.

Nicolo 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 *Neural Information Processing Systems (NeurIPS)*, 2020.

Thiparat Chotibut, Fryderyk Falniowski, Michał Misiurewicz, and Georgios Piliouras. The route to chaos in routing games: When is price of anarchy too optimistic? *Advances in Neural Information Processing Systems*, 33:766–777, 2020.

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

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 *Neural Information Processing Systems (NeurIPS)*, 2022.

Jonas Katona, Xiuyuan Wang, and Andre Wibisono. A symplectic analysis of alternating mirror descent. *arXiv preprint arXiv:2405.03472*, 2024.

Robert Kleinberg, Georgios Piliouras, and Éva Tardos. Multiplicative updates outperform generic no-regret learning in congestion games. In *ACM Symposium on Theory of Computing (STOC)*, 2009.

T. Y. Li and J. A. Yorke. Period three implies chaos. *Amer. Math. Monthly*, 82:985–992, 1975.

Panayotis Mertikopoulos, Christos Papadimitriou, and Georgios Piliouras. Cycles in adversarial regularized learning. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2703–2717. SIAM, 2018.

Lawrence Perko. *Differential equations and dynamical systems*, volume 7. Springer Science & Business Media, 2013.

Georgios Piliouras and Jeff S Shamma. Optimization despite chaos: Convex relaxations to complex limit sets via poincaré recurrence. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 861–873. SIAM, 2014.

Georgios Piliouras and Fang-Yi Yu. Multi-agent performative prediction: From global stability and optimality to chaos. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 1047–1074, 2023.

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

Ashkan Soleymani, Georgios Piliouras, and Gabriele Farina. Faster rates for no-regret learning in general games via cautious optimism. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 518–529, 2025.

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

Andre Wibisono, Molei Tao, and Georgios Piliouras. Alternating mirror descent for constrained min-max games. *Advances in Neural Information Processing Systems*, 35:35201–35212, 2022.