Fast Last-Iterate Convergence of Learning in Games Requires Forgetful Algorithms

Julien Grand-Clément (HEC Paris)

.

Joint with: Yang Cai, Gabriele Farina, Christian Kroer, Chung-Wei Lee, Haipeng Luo, Weiqiang Zhang

Workshop on Learning in Games - Toulouse 2024

This Talk

Main objective:
Understanding the last-iterate behavior of learning algorithms

Why it's interesting?
Recent applications: Poker[Brown and Sandholm, 2018], Stratego [Perolat et al., 2022], LLMs [Munos et al., 2023]...

Main result:
Optimistic FTRL does not admit last-iterate convergence rate that depends "nicely" on the players' dimensions and the payoff matrix

Regret minimization

Let $R : \Delta^d \to \mathbb{R}$ be 1-strongly convex.
Bregman divergence:

$$D_R(x, x') = R(x) - R(x') - \left\langle \nabla R(x'), x - x' \right\rangle.$$

Online Mirror Descent:

$$x^t = \operatorname*{argmin}_{x \in \Delta^d} \{ \left\langle \ell^{t-1}, x \right\rangle + \frac{1}{\eta} D_R(x, x^{t-1}) \} \qquad \text{(OMD)}$$

Follow-The-Regularized-Leader:

$$x^t = \operatorname*{argmin}_{x \in \Delta^d} \{ \langle \sum_{k=1}^{t-1} \ell^k, x \rangle + \frac{1}{\eta} R(x) \} \qquad \text{(FTRL)}$$

Note: OMD and FTRL are the same for Legendre regularizers.

Optimistic algorithms [Rakhlin and Sridharan, 2013, Syrgkanis et al., 2015]

Optimistic Online Mirror Descent:

$$\begin{aligned} \widehat{x}^t &= \operatorname*{argmin}_{x \in \Delta^d} \{ \langle \ell^{t-1}, x \rangle + \frac{1}{\eta} D_R(x, \widehat{x}^{t-1}) \} \\ x^t &= \operatorname*{argmin}_{x \in \Delta^d} \{ \langle \ell^{t-1}, x \rangle + \frac{1}{\eta} D_R(x, \widehat{x}^{t}) \} \end{aligned} \qquad \text{(OOMD)}$$

Optimistic Follow-The-Regularized-Leader:

$$x^t = \operatorname*{argmin}_{x \in \Delta^d} \left\{ \langle \sum_{k=1}^{t-1} \ell^k + \ell^{t-1}, x \rangle + \frac{1}{\eta} R(x) \right\} \qquad \text{(OFTRL)}$$

Optimistic algorithms [Rakhlin and Sridharan, 2013, Syrgkanis et al., 2015]

Optimistic Online Mirror Descent:

$$\widehat{x}^t = \operatorname*{argmin}_{x \in \Delta^d} \{\langle \ell^{t-1}, x \rangle + \frac{1}{\eta} D_R(x, \widehat{x}^{t-1})\}$$
$$x^t = \operatorname*{argmin}_{x \in \Delta^d} \{\langle \ell^{t-1}, x \rangle + \frac{1}{\eta} D_R(x, \widehat{x}^{t})\} \tag{OOMD}$$

Optimistic Follow-The-Regularized-Leader:

$$x^t = \operatorname*{argmin}_{x \in \Delta^d} \left\{ \langle \sum_{k=1}^{t-1} \ell^k + \ell^{t-1}, x \rangle + \frac{1}{\eta} R(x) \right\} \tag{OFTRL}$$

Two important algorithms:

- OGD: OOMD with $R = \frac{1}{2}\|\cdot\|_2^2$
- OMWU: OFTRL/OOMD with $R =$ negative entropy

Matrix games

Optimization problem:

$$\min_{x \in \Delta^{d_1}} \max_{y \in \Delta^{d_2}} x^\top A y$$

Goal: Compute $(x^\star, y^\star)$ with $\text{DualityGap}(x^\star, y^\star) = 0$, where

$$\text{DualityGap}(x^\star, y^\star) := \max_{y \in \Delta^{d_2}} (x^\star)^\top A y - \min_{x \in \Delta^{d_1}} x^\top A y^\star.$$

Self-play: both x-player and y-player use regret minimizers, using

$$\ell_x^t = A y^t, \ell_y^t = -A^\top x^t.$$

Matrix games

Optimization problem:

$$\min_{x \in \Delta^{d_1}} \max_{y \in \Delta^{d_2}} x^\top A y$$

Goal: Compute $(x^\star, y^\star)$ with $\text{DualityGap}(x^\star, y^\star) = 0$, where

$$\text{DualityGap}(x^\star, y^\star) := \max_{y \in \Delta^{d_2}} (x^\star)^\top A y - \min_{x \in \Delta^{d_1}} x^\top A y^\star.$$

Self-play: both x-player and y-player use regret minimizers, using

$$\ell_x^t = A y^t, \ell_y^t = -A^\top x^t.$$

Two important algorithms:

- OGDA: both players use OGD
- "OMWU": both players use OMWU

Advantages of OMWU over OGDA:

- Logarithmic dependence on the size of payoff matrix
- Closed-form updates:

$$x^t[i] \ \propto \ x^1[i] \times \exp\left(-\eta\left(\sum_{k=1}^{t-1} \ell^k[i] + \ell^{t-1}[i]\right)\right).$$

- $\tilde{O}(1/T)$ ergodic convergence to (coarse) correlated equilibrium in general-sum games [Daskalakis et al., 2021, Anagnostides et al., 2022]

Notions of convergence

Consider the sequences $\{x^t\}, \{y^t\}$ computed by self-play.

We have **ergodic convergence** when

$$\lim_{T \to +\infty} \text{DualityGap}(\bar{x}^T, \bar{y}^T) = 0.$$

[Rakhlin and Sridharan, 2013, Syrgkanis et al., 2015]: $O\left(\frac{1}{T}\right)$ ergodic convergence rate for OGDA/OMWU.

Notions of convergence

Consider the sequences $\{x^t\}, \{y^t\}$ computed by self-play.

We have **ergodic convergence** when

$$\lim_{T \to +\infty} \text{DualityGap}(\bar{x}^T, \bar{y}^T) = 0.$$

[Rakhlin and Sridharan, 2013, Syrgkanis et al., 2015]: $O\left(\frac{1}{T}\right)$ ergodic convergence rate for OGDA/OMWU.

We have **last-iterate convergence** when

$$\lim_{T \to +\infty} \text{DualityGap}(x^T, y^T) = 0.$$

Advantage: less computation (no need to average)

Convergence of optimistic algorithms

Last-iterate dynamics of OGDA:

- Unconstrained setting [Daskalakis et al., 2018, Hsieh et al., 2019, Liang and Stokes, 2019, Golowich et al., 2020]
- Matrix games: linear convergence with metric subregularity constants [Wei et al., 2021]
- Matrix games: convergence in $O(1/\sqrt{T})$ [Cai et al., 2022, Gorbunov et al., 2022]

Convergence of optimistic algorithms

Last-iterate dynamics of OGDA:

- Unconstrained setting [Daskalakis et al., 2018, Hsieh et al., 2019, Liang and Stokes, 2019, Golowich et al., 2020]
- Matrix games: linear convergence with metric subregularity constants [Wei et al., 2021]
- Matrix games: convergence in $O(1/\sqrt{T})$ [Cai et al., 2022, Gorbunov et al., 2022]

Last-iterate dynamics of OMWU:

- Unique N.E.: convergence with fixed step sizes (could be exponentially small), no rate [Daskalakis and Panageas, 2019]
- Unique N.E.: linear convergence with metric subregularity constants [Wei et al., 2021]
- Convergence for adaptive step sizes, no rate [Mertikopoulos et al., 2019, Hsieh et al., 2021]

Convergence of optimistic algorithms

Last-iterate dynamics of OGDA:

- Unconstrained setting [Daskalakis et al., 2018, Hsieh et al., 2019, Liang and Stokes, 2019, Golowich et al., 2020]
- Matrix games: linear convergence with metric subregularity constants [Wei et al., 2021]
- Matrix games: convergence in $O(1/\sqrt{T})$ [Cai et al., 2022, Gorbunov et al., 2022]

Last-iterate dynamics of OMWU:

- Unique N.E.: convergence with fixed step sizes (could be exponentially small), no rate [Daskalakis and Panageas, 2019]
- Unique N.E.: linear convergence with metric subregularity constants [Wei et al., 2021]
- Convergence for adaptive step sizes, no rate [Mertikopoulos et al., 2019, Hsieh et al., 2021]
- This work: rate of convergence of OMWU for matrix games?

Convergence of optimistic algorithms

For matrix games:

Algorithm	Ergodic	Last-iterate	Last-iterate (trick)
OGDA	$\frac{\text{poly}(d_1,d_2)L_2}{T}$	$\frac{\text{poly}(d_1,d_2)L_2}{\sqrt{T}}$	$\frac{C_1}{(1+C_2)^T}$
OMWU	$\frac{\text{polylog}(d_1,d_2)L_1}{T}$	o(1); rate?	$\frac{C_3}{(1+C_4)^T}$ (unique N.E.)

Convergence of optimistic algorithms

For matrix games:

Algorithm	Ergodic	Last-iterate	Last-iterate (trick)
OGDA	$\frac{\text{poly}(d_1,d_2)L_2}{T}$	$\frac{\text{poly}(d_1,d_2)L_2}{\sqrt{T}}$	$\frac{C_1}{(1+C_2)^T}$
OMWU	$\frac{\text{polylog}(d_1,d_2)L_1}{T}$	o(1); rate?	$\frac{C_3}{(1+C_4)^T}$ (unique N.E.)

Note: C_1, C_2, C_3, C_4 obtained from metric subregularity:

$$\text{DualityGap}(x, y) \geq c \cdot \text{dist}((x, y), \text{ set of N.E.})$$

$\Rightarrow$ may be arbitrarily bad even with fixed $d_1, d_2, \max_{i,j} |A_{ij}|$.

Our main theorem

Theorem (Informal)

Consider two-player zero-sum games with matrix entries in $[0, 1]$*, and* d_1 *and* d_2 *are the number of actions.*
For OMWU with constant step size, no function f *can satisfy*

1. $\mathrm{DualityGap}(x^T, y^T) \le f(d_1, d_2, T)$ *for all* T*.*
2. $\lim_{T\to\infty} f(d_1, d_2, T) \to 0$*.*

Our main theorem

Theorem (Informal)

Consider two-player zero-sum games with matrix entries in $[0, 1]$*, and* d_1 *and* d_2 *are the number of actions.*
For OMWU with constant step size, no function f *can satisfy*

1. $\text{DualityGap}(x^T, y^T) \leq f(d_1, d_2, T)$ *for all* T*.*
2. $\lim_{T\to\infty} f(d_1, d_2, T) \to 0$*.*

Holds for OFTRL with regularizer = entropy/log/squared L_2 norm/Tsallis entropy

Our main theorem

Theorem (Informal)

Consider two-player zero-sum games with matrix entries in $[0, 1]$*, and* d_1 *and* d_2 *are the number of actions.*
For OMWU with constant step size, no function f *can satisfy*

1. $\text{DualityGap}(x^T, y^T) \leq f(d_1, d_2, T)$ *for all* T*.*
2. $\lim_{T\to\infty} f(d_1, d_2, T) \to 0.$

Holds for OFTRL with regularizer = entropy/log/squared L_2 norm/Tsallis entropy

Idea of the proof:

- We construct a 2x2 matrix game A_δ parametrized by $\delta > 0$.
- After $\Omega(1/\eta\delta)$ iterations of OFTRL, the duality gap is a constant c, a universal constant depending on the regularizer.

A difficult matrix game for OFTRL

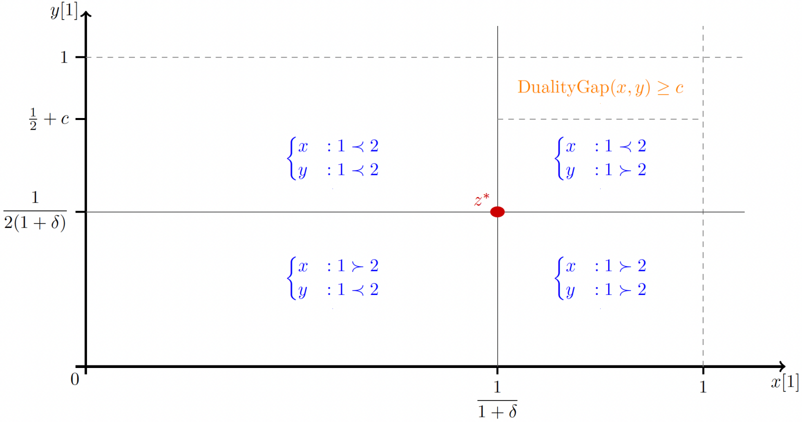
Consider the matrix game A_δ with $\delta \in (0, 1/2]$:

$$A_\delta := \begin{bmatrix} \frac{1}{2} + \delta & \frac{1}{2} \\ 0 & 1 \end{bmatrix}.$$

A_δ has a unique Nash equilibrium: $x^*[1] = \frac{1}{1+\delta}, y^*[1] = \frac{1}{2(1+\delta)}$.

Bad region: $x[1] \geq \frac{1}{1+\delta}, y[1] \geq \frac{1}{2} + c \Rightarrow \text{DualityGap}(x, y) \geq c$.

A difficult matrix game for OFTRL



Reformulating OFTRL

2x2 game $\Rightarrow$ we focus on $x[1], y[1]$. Define

$$F_{\eta,R}(e) := \operatorname*{argmin}_{x\in[0,1]} \left\{ x \cdot e + \frac{1}{\eta} R(x) \right\}$$
$$e_x^t := \ell_x^t[1] - \ell_x^t[2]$$
$$E_x^t := \sum_{k=1}^{t} e_x^k$$

We can rewrite OFTRL:

$$x^t[1] = F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right)$$
$$y^t[1] = F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right) \qquad \text{(OFTRL)}$$

Assumptions on the regularizers

Define

$$F_{\eta,R}(e) := \underset{x\in[0,1]}{\operatorname{argmin}}\{x \cdot e + \frac{1}{\eta}R(x)\}.$$

Important: $F_{\eta,R} : \mathbb{R} \to [0,1]$ is non-increasing.

Assumption

We assume that the regularizer R satisfies the following properties: the function $F_{\eta,R} : \mathbb{R} \to [0,1]$ is

1. **Unbiased:** $F_{\eta,R}(0) = \frac{1}{2}$.
2. **Rational:** $\lim_{E\to-\infty} F_{\eta,R}(E) = 1$ *and* $\lim_{E\to+\infty} F_{\eta,R}(E) = 0$.
3. **Lipschitz continuous:** *There exists $L \geq 0$ such that $F_{1,R}$ is L-Lipschitz.*

More assumptions here: (1).

Reformulating OFTRL

$$x^t[1] = F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right)$$
$$y^t[1] = F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right) \qquad \text{(OFTRL)}$$

Important note: $e_y^t \in [-\delta, 1]$
$\Rightarrow$ if E_y^t is large, it takes $\Omega(1/\delta)$ iterations to make it close to 0.

Numerical experiments

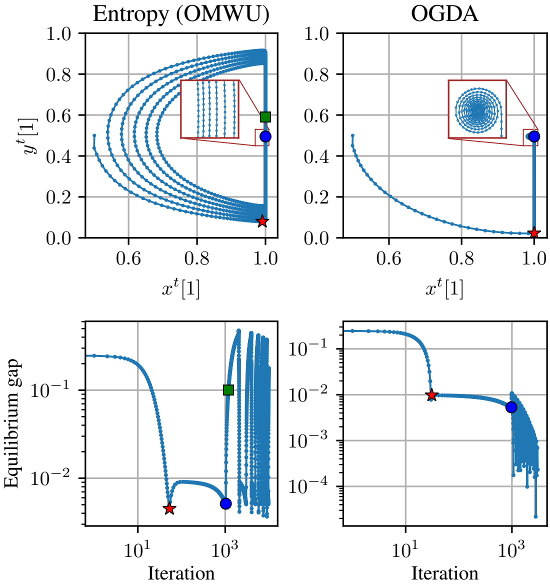


Figure: Dynamics produced by OMWU and OGDA in the same game A_δ.

Numerical experiments

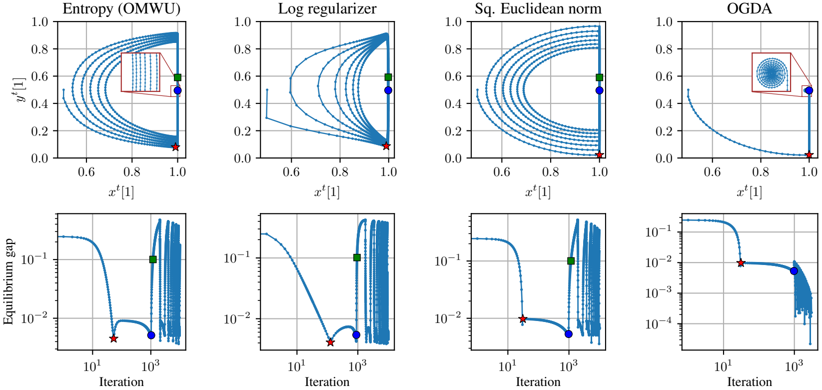


Figure: Dynamics produced by variants of OFTRL with different regularizers and OGDA in the same game A_δ.

$$x^t[1] = F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right)$$
$$y^t[1] = F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right) \qquad \text{(OFTRL)}$$

Numerical experiments

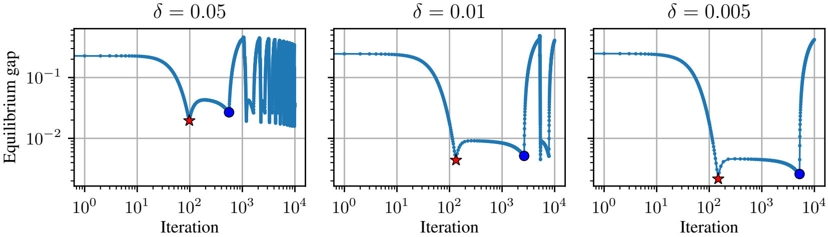
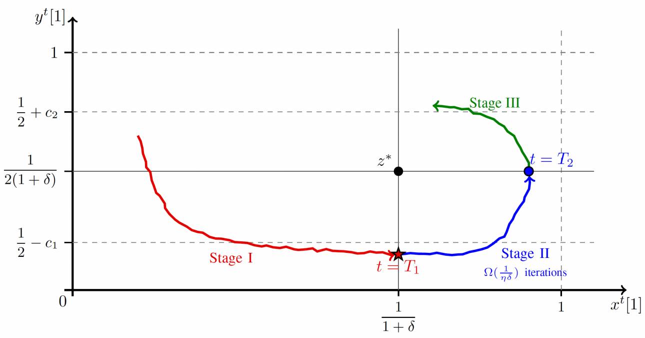


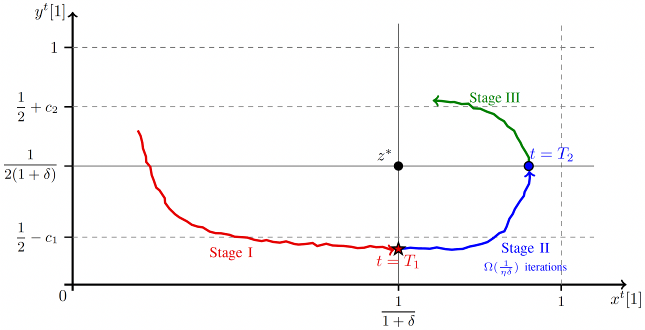
Figure: Influence of $\delta > 0$ on the duality gaps of OMWU after 10^4 iterations.



$$x^t[1] = F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right)$$
$$y^t[1] = F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right) \qquad \text{(OFTRL)}$$

Stage I: Starting at $x^1[1] = y^1[1] = 1/2$, we prove

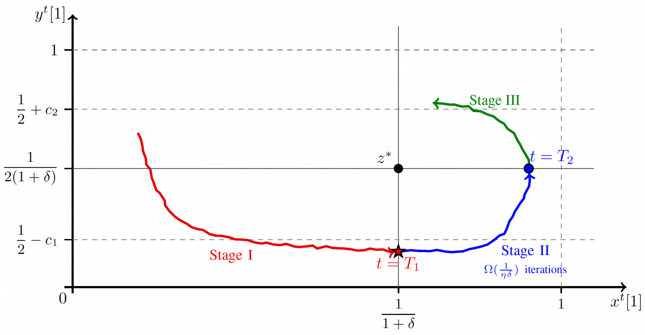
- $x^t[1]$ increases until T_1 s.t. $x^{T_1}[1] \geq \frac{1}{1+\delta}$.



$$x^t[1] = F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right)$$
$$y^t[1] = F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right) \quad \text{(OFTRL)}$$

Stage I: Starting at $x^1[1] = y^1[1] = 1/2$, we prove

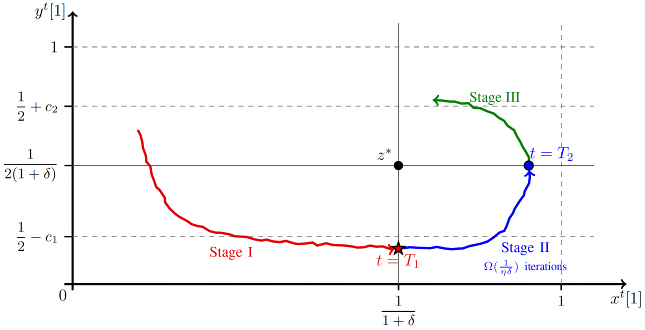
- $x^t[1]$ increases until T_1 s.t. $x^{T_1}[1] \geq \frac{1}{1+\delta}$.
- $e_y^t \geq 0$: action 1 $\prec$ action 2 for the y-player, $y^t[1]$ decreases.



$$x^t[1] = F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right)$$
$$y^t[1] = F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right) \qquad \text{(OFTRL)}$$

Stage I: Starting at $x^1[1] = y^1[1] = 1/2$, we prove

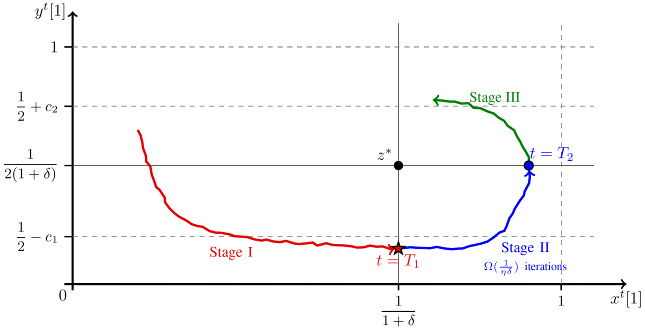
- $x^t[1]$ increases until T_1 s.t. $x^{T_1}[1] \geq \frac{1}{1+\delta}$.
- $e_y^t \geq 0$: action 1 $\prec$ action 2 for the y-player, $y^t[1]$ decreases.
- At the last period T_1, we have $y^{T_1}[1] \leq \frac{1}{2} - c_1$.



$$x^t[1] = F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right)$$
$$y^t[1] = F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right) \qquad \text{(OFTRL)}$$

Stage II:

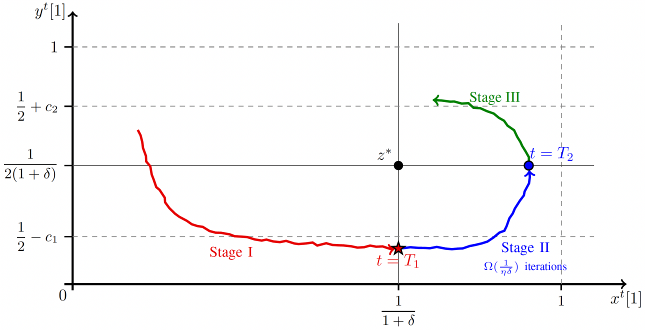
- $-\delta \leq e_y^t < 0 \Rightarrow y^t[1]$ increases, at most by $\eta L\delta$.



$$x^t[1] = F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right)$$
$$y^t[1] = F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right) \qquad \text{(OFTRL)}$$

Stage II:

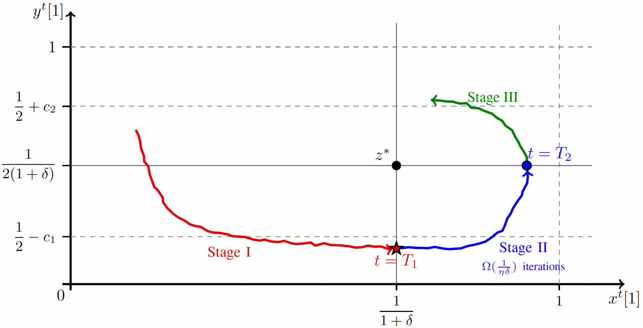
- $-\delta \leq e_y^t < 0 \Rightarrow y^t[1]$ increases, at most by $\eta L\delta$.
- Stage II lasts until T_2 s.t. $y^{T_2}[1] \geq 1/2(1+\delta)$.
- Thus $T_2 - T_1 = \Omega(c_1/\eta L\delta)$.



$$
\begin{aligned}
x^t[1] &= F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right) \\
y^t[1] &= F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right)
\end{aligned} \qquad \text{(OFTRL)}
$$

Stage II:

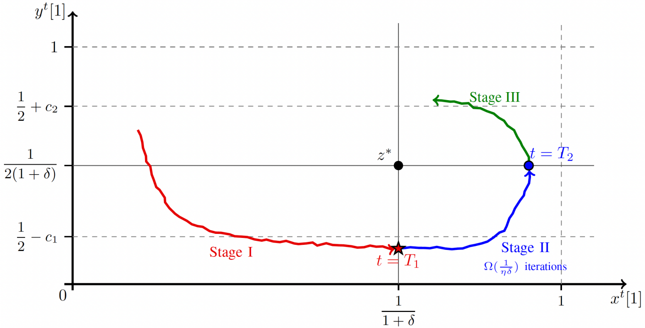
- $-\delta \leq e_y^t < 0 \Rightarrow y^t[1]$ increases, at most by $\eta L \delta$.
- Stage II lasts until T_2 s.t. $y^{T_2}[1] \geq 1/2(1+\delta)$.
- Thus $T_2 - T_1 = \Omega(c_1/\eta L \delta)$.
- $e_x^t < 0$, and $x^t[1]$ keeps growing closer to 1: $E_x^{T_2} \leq E_x^{T_1} - \Omega(1/\eta L \delta)$.



$$x^t[1] = F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right)$$
$$y^t[1] = F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right) \qquad \text{(OFTRL)}$$

Stage III:

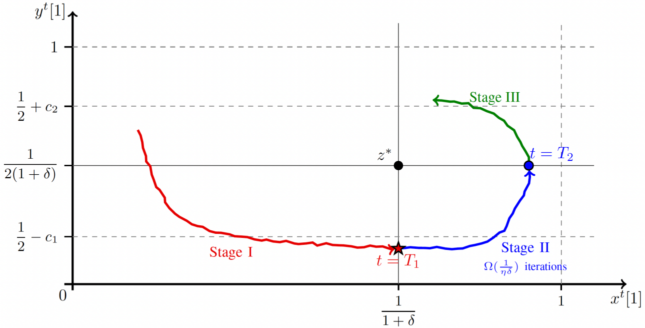
- $y^t[1]$ keeps increasing until $x^t[1] \leq 1/(1+\delta)$.



$$x^t[1] = F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right)$$
$$y^t[1] = F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right) \qquad \text{(OFTRL)}$$

Stage III:

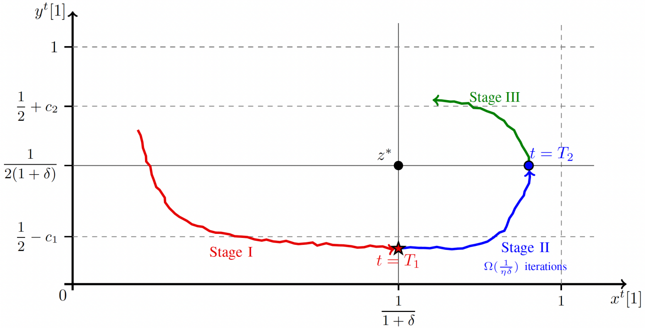
- $y^t[1]$ keeps increasing until $x^t[1] \le 1/(1+\delta)$.
- $x^t[1] \le 1/(1+\delta) \Rightarrow E_x^{T_1} \le E_x^t$.



$$x^t[1] = F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right)$$
$$y^t[1] = F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right) \qquad \text{(OFTRL)}$$

Stage III:

- $y^t[1]$ keeps increasing until $x^t[1] \leq 1/(1+\delta)$.
- $x^t[1] \leq 1/(1+\delta) \Rightarrow E_x^{T_1} \leq E_x^t$.
- But at T_2, we have $E_x^{T_2} \leq E_x^{T_1} - \Omega(1/\eta L\delta)$. Since $e_x^t \leq 1$, we still have $x^t[1] \geq 1/(1+\delta)$ after $\Omega(1/\eta L\delta)$ steps.



$$x^t[1] = F_{\eta,R}\left(E_x^{t-1} + e_x^{t-1}\right)$$
$$y^t[1] = F_{\eta,R}\left(E_y^{t-1} + e_y^{t-1}\right) \qquad \text{(OFTRL)}$$

Stage III:

- $y^t[1]$ keeps increasing until $x^t[1] \le 1/(1+\delta)$.
- $x^t[1] \le 1/(1+\delta) \Rightarrow E_x^{T_1} \le E_x^t$.
- But at T_2, we have $E_x^{T_2} \le E_x^{T_1} - \Omega(1/\eta L\delta)$. Since $e_x^t \le 1$, we still have $x^t[1] \ge 1/(1+\delta)$ after $\Omega(1/\eta L\delta)$ steps.
- Until $T_3 = T_2 + \Omega(1/\eta L\delta)$, $e_y^t \le 0$, and $y^{T_3}[1] \ge \frac{1}{2} + c_2$.

Main theorem

We proved:
There is a universal constant c (dependent on the regularizer) such that for any $\delta > 0$*, we can find a game* A_δ *where the duality gap for OFTRL is at least c after* $\Omega(1/\delta)$ *rounds.*

Main theorem

We proved:
There is a universal constant c (dependent on the regularizer) such that for any $\delta > 0$, we can find a game A_δ where the duality gap for OFTRL is at least c after $\Omega(1/\delta)$ rounds.

Formal statement:

Theorem

Assume the regularizer R satisfies our assumptions with universal constant $c_1, c_2, , \hat{\delta}, L > 0$. Let $\delta \in (0, \hat{\delta})$
The OFTRL dynamics on A_δ with any step size $\eta \leq \frac{1}{4L}$ satisfies the following: there exists an iteration $t \geq \frac{c_1}{3\eta L \delta}$ with a duality gap of at least c_2.

Conclusion

- **Main result**: Negative result for the rate of convergence of OMWU (OFTRL).

- **Next steps**:
  Universal best-iterate convergence rate?
  Rate for adaptive step sizes?
  Mixing OGD/OMWU?

- Preprint: `https://arxiv.org/abs/2406.10631`
  Slides/code on my website

Conclusion

- **Main result**: Negative result for the rate of convergence of OMWU (OFTRL).

- **Next steps**:
  Universal best-iterate convergence rate?
  Rate for adaptive step sizes?
  Mixing OGD/OMWU?

- Preprint: `https://arxiv.org/abs/2406.10631`
  Slides/code on my website

Thank you!

References I

Anagnostides, I., Daskalakis, C., Farina, G., Fishelson, M., Golowich, N., and Sandholm, T. (2022).
Near-optimal no-regret learning for correlated equilibria in multi-player general-sum games.
In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*.

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

Cai, Y., Oikonomou, A., and Zheng, W. (2022).
Finite-time last-iterate convergence for learning in multi-player games.
In *Advances in Neural Information Processing Systems (NeurIPS)*.

References II

Daskalakis, C., Fishelson, M., and Golowich, N. (2021).
Near-optimal no-regret learning in general games.
Advances in Neural Information Processing Systems (NeurIPS).

Daskalakis, C., Ilyas, A., Syrgkanis, V., and Zeng, H. (2018).
Training gans with optimism.
In *International Conference on Learning Representations (ICLR).*

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

References III

Duchi, J., Hazan, E., and Singer, Y. (2011).
Adaptive subgradient methods for online learning and stochastic optimization.
Journal of machine learning research, 12(7).

Golowich, N., Pattathil, S., and Daskalakis, C. (2020).
Tight last-iterate convergence rates for no-regret learning in multi-player games.
Advances in neural information processing systems (NeurIPS).

Gorbunov, E., Taylor, A., and Gidel, G. (2022).
Last-iterate convergence of optimistic gradient method for monotone variational inequalities.
In *Advances in Neural Information Processing Systems*.

References IV

Hsieh, Y.-G., Antonakopoulos, K., and Mertikopoulos, P. (2021).
Adaptive learning in continuous games: Optimal regret bounds and convergence to nash equilibrium.
In *Conference on Learning Theory*, pages 2388–2422. PMLR.

Hsieh, Y.-G., Iutzeler, F., Malick, J., and Mertikopoulos, P. (2019).
On the convergence of single-call stochastic extra-gradient methods.
Advances in Neural Information Processing Systems, 32.

Liang, T. and Stokes, J. (2019).
Interaction matters: A note on non-asymptotic local convergence of generative adversarial networks.
In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 907–915. PMLR.

References V

Mertikopoulos, P., Lecouat, B., Zenati, H., Foo, C.-S., Chandrasekhar, V., and Piliouras, G. (2019).
Optimistic mirror descent in saddle-point problems: Going the extra (gradient) mile.
In *International Conference on Learning Representations (ICLR).*

Munos, R., Valko, M., Calandriello, D., Azar, M. G., Rowland, M., Guo, Z. D., Tang, Y., Geist, M., Mesnard, T., Michi, A., et al. (2023).
Nash learning from human feedback.
arXiv preprint arXiv:2312.00886.

References VI

Perolat, J., De Vylder, B., Hennes, D., Tarassov, E., Strub, F., de Boer, V., Muller, P., Connor, J. T., Burch, N., Anthony, T., et al. (2022).
Mastering the game of stratego with model-free multiagent reinforcement learning.
Science, 378(6623):990–996.

Rakhlin, S. and Sridharan, K. (2013).
Optimization, learning, and games with predictable sequences.
Advances in Neural Information Processing Systems.

Syrgkanis, V., Agarwal, A., Luo, H., and Schapire, R. E. (2015).
Fast convergence of regularized learning in games.
Advances in Neural Information Processing Systems (NeurIPS).

References VII

Wei, C.-Y., Lee, C.-W., Zhang, M., and Luo, H. (2021).
Linear last-iterate convergence in constrained saddle-point optimization.
In *International Conference on Learning Representations (ICLR)*.

Assumptions on the regularizers

Define

$$F_{\eta,R}(e) := \underset{x\in[0,1]}{\operatorname{argmin}}\{x \cdot e + \frac{1}{\eta} R(x)\}.$$

Assumption (Informal)

There are some constants such that

$$F_{1,R}(E) \geq \frac{1}{1+\delta} \Rightarrow F_{1,R}\left(E - \Omega\left(\frac{1}{\delta}\right)\right) \geq F_{1,R}(E) + \Omega(\delta) \tag{1}$$

$$F_{1,R}(E) \geq \frac{1}{2(1+\delta)} \Rightarrow F_{1,R}\left(E - \Omega(\delta)\right) \geq \frac{1}{2} + \Omega(1) \tag{2}$$

Both our assumption hold for the negative entropy, squared Euclidean norm, the log barrier, and the Tsallis entropy regularizers. Link to main presentation: (2).

Assumptions on the regularizers

Define

$$F_{\eta,R}(e) := \operatorname*{argmin}_{x\in[0,1]}\{x \cdot e + \frac{1}{\eta}R(x)\}.$$

Assumption

Let L be the Lipschitness constant of $F_{1,R}$. Denote constant $c_1 = \frac{1}{2} - F_{1,R}(\frac{1}{20L})$. There exist universal constants $\delta', c_2 > 0$ and $c_3 \in (0, \frac{1}{2}]$ such that for any $0 < \delta \leq \delta'$,

1. *If $F_{1,R}(E) \geq \frac{1}{1+\delta}$, then $F_{1,R}(-\frac{c_1^2}{30L\delta} + E) \geq \frac{1+c_3}{1+c_3+\delta}$*
2. *If $F_{1,R}(E) \geq \frac{1}{2(1+\delta)}$, then $F_{1,R}(-\frac{c_3 c_1^2}{120L} + \frac{\delta}{4L} + E) \geq \frac{1}{2} + c_2$.*

Both our assumption hold for the negative entropy, squared Euclidean norm, the log barrier, and the negative Tsallis entropy regularizers. Link to main presentation: (2)

Other convergence rates

Following [Wei et al., 2021]:

Corollary

Let $\delta \in (0, \frac{1}{2})$. For OMWU with step size $\eta \leq \frac{1}{8}$ on A_δ satisfies

$$\text{DualityGap}(x^T, y^T) \leq \frac{1200e^{\frac{10}{\delta}}}{\eta} \cdot \frac{1}{\sqrt{T}}, \forall\ T \geq 1.$$

$\Rightarrow$ Problem-constant independent best-iterate rate for OMWU:

Theorem

Let $\delta \in (0, \frac{1}{32})$. For OMWU with step size $\eta \leq \frac{1}{8}$

$$\min_{t \in [T]} \text{DualityGap}(x^t, y^t) \leq O\left(\frac{1}{\eta \ln T}\right), \forall\ T \geq 2.$$

Adaptive stepsizes

Adaptive stepsize [Duchi et al., 2011]: $\eta_t = 1/\sqrt{\epsilon + \sum_{k=1}^{t-1} \|\boldsymbol{\ell}_k\|_k^2}$

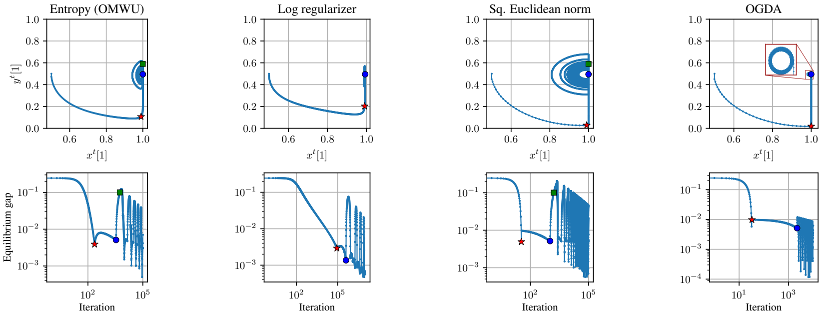


Figure: Here $\delta := 10^{-2}$ and adaptive step size with $\epsilon = 0.1$.

Mixing OGD and OMWU 1/2

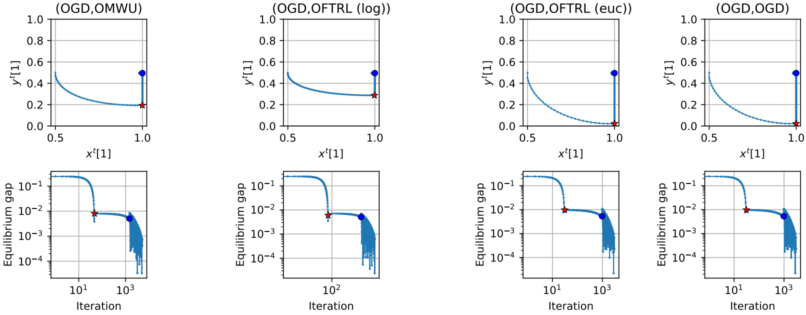


Figure: Dynamics and duality gap when the x-player uses OGD. We choose $\delta = 0.01$ and $\eta = 0.1$ in all figures.

Mixing OGD and OMWU 2/2

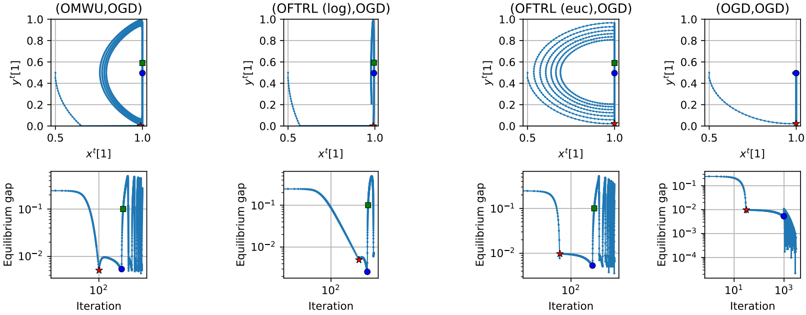


Figure: Dynamics and duality gap when the y-player uses OGD. We choose $\delta = 0.01$ and $\eta = 0.1$ in all figures.