# Appendix for "Let's be Honest: An Optimal No-Regret Framework for Zero-Sum Games"

**Ehsan Asadi Kangarshahi** [* 1] **Ya-Ping Hsieh** [* 1] **Mehmet Fatih Sahin** [1] **Volkan Cevher** [1]

## A. Equivalence Formulations of Optimistic Mirror Descent

In this appendix, we show that the $\mathbf{x}_t$ iterates in (2) of the main text is equivalent to the following iterates given in (Chiang et al., 2012; Rakhlin & Sridharan, 2013):

$$\begin{cases} \mathbf{x}_t & = MD_\eta\left(\tilde{\mathbf{x}}_t, -A\mathbf{y}_{t-1}\right) \\ \tilde{\mathbf{x}}_{t+1} & = MD_\eta\left(\tilde{\mathbf{x}}_t, -A\mathbf{y}_t\right) \end{cases}. \tag{A.1}$$

By the optimality condition for (A.1), we have

$$\nabla\psi(\mathbf{x}_t) = \nabla\psi(\tilde{\mathbf{x}}_t) - \eta\left(-A\mathbf{y}_{t-1}\right), \tag{A.2}$$

$$\nabla\psi(\tilde{\mathbf{x}}_t) = \nabla\psi(\tilde{\mathbf{x}}_{t-1}) - \eta\left(-A\mathbf{y}_{t-1}\right), \tag{A.3}$$

$$\nabla\psi(\tilde{\mathbf{x}}_{t-1}) = \nabla\psi(\mathbf{x}_{t-1}) + \eta\left(-A\mathbf{y}_{t-2}\right). \tag{A.4}$$

We hence get (2) by applying (A.4) to (A.3) and then (A.3) to (A.2).

## B. Optimistic Mirror Descent

In this appendix, we prove **Theorem 2**, restated below for convenience.

**Theorem 1.** *Suppose two players of a zero-sum game have played $T$ rounds according to **Algorithm 1** and **2** with $\eta = \frac{1}{2|A|_{\max}}$. Then*

1. *The $\mathbf{x}$-player suffers a $O\left(\frac{\log(T)}{T}\right)$ regret:*

$$\max_{\mathbf{z}\in\Delta_m} \sum_{t=3}^{T} \langle \mathbf{z}_t - \mathbf{z}, -A\mathbf{w}_t\rangle \leq \Big(\log(T-2)+1\Big)\Big(20+\log m+\log n\Big)|A|_{\max} \\ = O\left(\log T\right) \tag{B.1}$$

   *and similarly for the $\mathbf{y}$-player.*

2. *The strategies $(\mathbf{z}_T, \mathbf{w}_T)$ constitutes an $O\left(\frac{1}{T}\right)$-approximate equilibrium to the value of the game:*

$$|V - \langle \mathbf{z}_T, A\mathbf{w}_T\rangle| \leq \frac{\Big(20+\log m+\log n\Big)|A|_{\max}}{T-2} = O\left(\frac{1}{T}\right). \tag{B.2}$$

[*]Equal contribution [1]LIONS, EPFL, Switzerland. Correspondence to: Ya-Ping Hsieh <ya-ping.hsieh@epfl.ch>, Volkan Cevher <volkan.cevher@epfl.ch>.

*Proceedings of the $35^{th}$ International Conference on Machine Learning*, Stockholm, Sweden, PMLR 80, 2018. 

*Proof.* Define $\mathbf{x}^*$ as

$$\mathbf{x}^* = \arg\min_{\mathbf{x}\in\Delta_m} \left\langle \mathbf{x}, -A\left(\frac{1}{T-2}\sum_{t=3}^{T}\mathbf{y}_t\right)\right\rangle. \tag{B.3}$$

We define an auxiliary individual regret $\mathrm{R}_T^{\mathbf{x}}$ as

$$\mathrm{R}_T^{\mathbf{x}} \coloneqq \sum_{t=3}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, -A\mathbf{y}_t\rangle. \tag{B.4}$$

Notice that this is the regret on the $\mathbf{x}_t$ sequence versus $\mathbf{y}_t$ sequence, while we are playing $\mathbf{z}_t$'s and $\mathbf{w}_t$'s in the algorithm.

We then have

$$\begin{aligned}
\mathrm{R}_T^{\mathbf{x}} &= \sum_{t=3}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, -A\mathbf{y}_t\rangle \\
&= \langle \mathbf{x}_3 - \mathbf{x}^*, -A\mathbf{y}_3\rangle + \sum_{t=4}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, -A\mathbf{y}_t\rangle \\
&\le 2|A|_{\max} + \sum_{t=4}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, -A\mathbf{y}_t - \mathbf{g}_{t-1}\rangle + \sum_{t=4}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, \mathbf{g}_{t-1}\rangle
\end{aligned}$$

where $\mathbf{g}_t \coloneqq -2(t-2)A\mathbf{w}_t + 3(t-3)A\mathbf{w}_{t-1} - (t-4)A\mathbf{w}_{t-2}$. Inserting $\mathbf{w}_t = \frac{1}{t-2}\sum_{i=3}^{t}\mathbf{y}_i$ into the definition of $\mathbf{g}_t$, we get $\mathbf{g}_t = -2A\mathbf{y}_t + A\mathbf{y}_{t-1}$. Straightforward calculation then shows:

$$\begin{aligned}
\mathrm{R}_T^{\mathbf{x}} &\le 2|A|_{\max} + \sum_{t=4}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, -A\mathbf{y}_t + 2A\mathbf{y}_{t-1} - A\mathbf{y}_{t-2}\rangle + \sum_{t=4}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, -2A\mathbf{y}_{t-1} + A\mathbf{y}_{t-2}\rangle \\
&= 2|A|_{\max} + \sum_{t=4}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, (-A\mathbf{y}_t + A\mathbf{y}_{t-1}) - (-A\mathbf{y}_{t-1} + A\mathbf{y}_{t-2})\rangle \\
&\qquad + \frac{1}{\eta}\sum_{t=4}^{T}\Big(D(\mathbf{x}^*, \mathbf{x}_{t-1}) - D(\mathbf{x}^*, \mathbf{x}_t) - D(\mathbf{x}_t, \mathbf{x}_{t-1})\Big) \\
&= 2|A|_{\max} + \sum_{t=4}^{T-1}\langle \mathbf{x}_t - \mathbf{x}_{t+1}, -A\mathbf{y}_t + A\mathbf{y}_{t-1}\rangle + \langle \mathbf{x}_4 - \mathbf{x}^*, A\mathbf{y}_3 - A\mathbf{y}_2\rangle \\
&\qquad + \langle \mathbf{x}_T - \mathbf{x}^*, -A\mathbf{y}_T + A\mathbf{y}_{T-1}\rangle + \frac{1}{\eta}\sum_{t=4}^{T}\Big(D(\mathbf{x}^*, \mathbf{x}_{t-1}) - D(\mathbf{x}^*, \mathbf{x}_t) - D(\mathbf{x}_t, \mathbf{x}_{t-1})\Big) \\
&\le 10|A|_{\max} + \sum_{t=4}^{T-1}\langle \mathbf{x}_t - \mathbf{x}_{t+1}, -A\mathbf{y}_t + A\mathbf{y}_{t-1}\rangle \\
&\qquad + \frac{1}{\eta}\sum_{t=4}^{T}\Big(D(\mathbf{x}^*, \mathbf{x}_{t-1}) - D(\mathbf{x}^*, \mathbf{x}_t) - D(\mathbf{x}_t, \mathbf{x}_{t-1})\Big) \\
&\le 10|A|_{\max} + \sum_{t=4}^{T-1}\|\mathbf{x}_t - x_{t+1}\|_1 \cdot |A|_{\max} \cdot \|\mathbf{y}_t - \mathbf{y}_{t-1}\|_1 \\
&\qquad + \frac{1}{\eta}\Big(D(\mathbf{x}^*, \mathbf{x}_3) - D(\mathbf{x}^*, \mathbf{x}_T)\Big) + \sum_{t=4}^{T}\frac{-1}{\eta}D(\mathbf{x}_t, \mathbf{x}_{t-1}) \\
&\le 10|A|_{\max} + \frac{1}{2}\sum_{t=4}^{T-1}\Big(|A|_{\max}\cdot\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_1^2 + |A|_{\max}\cdot\|\mathbf{y}_t - \mathbf{y}_{t-1}\|_1^2\Big)
\end{aligned}$$

$$+ \frac{1}{\eta}\Big(D(\mathbf{x}^*, \mathbf{x}_3) - D(\mathbf{x}^*, \mathbf{x}_T)\Big) + \sum_{t=4}^{T} \frac{-1}{\eta} D(\mathbf{x}_t, \mathbf{x}_{t-1}).$$

Using the fact that $\psi$ is 1-strongly convex with respect to the $\ell_1$-norm, we have $-D(\mathbf{x}, \mathbf{x}') \le -\frac{1}{2}\|\mathbf{x} - \mathbf{x}'\|_1^2 \le 0$. Also, we have $D(\mathbf{x}^*, \mathbf{x}_3) \le \log m$. Combining these facts in the last inequality gives:

$$\begin{aligned}\mathrm{R}_T^{\mathbf{x}} \le 10|A|_{\max} + \frac{\log m}{\eta} + \frac{|A|_{\max}}{2} \sum_{t=4}^{T-1} \|\mathbf{x}_t - \mathbf{x}_{t+1}\|_1^2 \\ + \frac{|A|_{\max}}{2} \sum_{t=4}^{T-1} \|\mathbf{y}_t - \mathbf{y}_{t-1}\|_1^2 - \frac{1}{2\eta} \sum_{t=4}^{T} \|\mathbf{x}_{t-1} - \mathbf{x}_t\|_1^2.\end{aligned}$$

Similarly, for the second player we define

$$\mathrm{R}_T^{\mathbf{y}} \coloneqq \sum_{t=3}^{T} \big\langle \mathbf{y}_t - \mathbf{y}^*, A^\top \mathbf{x}_t \big\rangle \tag{B.5}$$

where $\mathbf{y}^* \coloneqq \arg\min_{\mathbf{y}} \left\langle \mathbf{y}, A^\top \left(\frac{1}{T-2}\sum_{t=3}^{T} \mathbf{x}_t\right)\right\rangle$. We then have

$$\begin{aligned}\mathrm{R}_T^{\mathbf{y}} \le 10|A|_{\max} + \frac{\log n}{\eta} + \frac{|A|_{\max}}{2} \sum_{t=4}^{T-1} \|\mathbf{y}_t - \mathbf{y}_{t+1}\|_1^2 \\ + \frac{|A|_{\max}}{2} \sum_{t=4}^{T-1} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_1^2 - \frac{1}{2\eta} \sum_{t=4}^{T} \|\mathbf{y}_{t-1} - \mathbf{y}_t\|_1^2.\end{aligned}$$

Setting $\eta = \frac{1}{2|A|_{\max}}$, we get

$$\mathrm{R}_T^{\mathbf{x}} + \mathrm{R}_T^{\mathbf{y}} \le \Big(20 + \log m + \log n\Big)|A|_{\max}. \tag{B.6}$$

Now, recalling that $\mathbf{z}_T = \frac{\sum_{t=3}^{T} \mathbf{x}_t}{T-2}$ and $\mathbf{w}_T = \frac{\sum_{t=3}^{T} \mathbf{y}_t}{T-2}$ and using the definition of $\mathrm{R}_T^{\mathbf{x}}$ and $\mathrm{R}_T^{\mathbf{y}}$, we get

$$\frac{1}{T-2}\Big(\mathrm{R}_T^{\mathbf{x}} + \mathrm{R}_T^{\mathbf{y}}\Big) = \max_{\mathbf{x}\in\Delta_m} \langle \mathbf{x}, A\mathbf{w}_T\rangle - \min_{\mathbf{y}\in\Delta_n} \langle \mathbf{z}_T, A\mathbf{y}\rangle. \tag{B.7}$$

Furthermore, by the definition of the value of the game, we have

$$\min_{\mathbf{y}\in\Delta_n} \langle \mathbf{z}_T, A\mathbf{y}\rangle \le V \le \max_{\mathbf{x}\in\Delta_m} \langle \mathbf{x}, A\mathbf{w}_T\rangle. \tag{B.8}$$

We also trivially have

$$\min_{\mathbf{y}\in\Delta_n} \langle \mathbf{z}_T, A\mathbf{y}\rangle \le \langle \mathbf{z}_T, A\mathbf{w}_T\rangle \le \max_{\mathbf{x}\in\Delta_m} \langle \mathbf{x}, A\mathbf{w}_T\rangle. \tag{B.9}$$

Combining (B.7) - (B.9) in (B.6) then establishes (4):

$$|V - \langle \mathbf{z}_T, A\mathbf{w}_T\rangle| \le \frac{\Big(20 + \log m + \log n\Big)|A|_{\max}}{T-2}.$$

We now turn to (3).

Let $\mathrm{R}_T^{\mathbf{z}} \coloneqq \max_{\mathbf{z}\in\Delta_m} \sum_{t=3}^{T} \langle \mathbf{z}_t - \mathbf{z}, -A\mathbf{w}_t\rangle$ and let $\tilde{\mathrm{R}}_T^{\mathbf{z}} \coloneqq \sum_{t=3}^{T} \langle \mathbf{z}_t - \mathbf{z}_t^*, -A\mathbf{w}_t\rangle$ where $\mathbf{z}_t^* = \arg\min_{\mathbf{z}\in\Delta_m} \langle \mathbf{z}, -A\mathbf{w}_t\rangle$. Evidently we have $\mathrm{R}_T^{\mathbf{z}} \le \tilde{\mathrm{R}}_T^{\mathbf{z}}$. Notice that (with $\mathbf{w}_t^*$ similarly defined)

$$\begin{aligned}\langle \mathbf{z}_t - \mathbf{z}_t^*, -A\mathbf{w}_t\rangle &= \langle \mathbf{z}_t^*, A\mathbf{w}_t\rangle - \langle \mathbf{z}_t, A\mathbf{w}_t\rangle \\ &\le \langle \mathbf{z}_t^*, A\mathbf{w}_t\rangle - \langle \mathbf{z}_t, A\mathbf{w}_t^*\rangle\end{aligned}$$

$$\leq \frac{\Big(20 + \log m + \log n\Big)|A|_{\max}}{t-2} \tag{B.10}$$

by (B.6) and (B.7). Using these inequalities, we get

$$\begin{aligned}
\frac{1}{T-2}\mathrm{R}_T^{\mathbf{z}} \leq \frac{1}{T-2}\tilde{\mathrm{R}}_T^{\mathbf{z}} &= \frac{1}{T-2}\sum_{t=3}^{T}\langle \mathbf{z}_t - \mathbf{z}_t^*, -A\mathbf{w}_t\rangle \\
&\leq \frac{1}{T-2}\sum_{t=3}^{T}\frac{\Big(20 + \log m + \log n\Big)|A|_{\max}}{t-2} \\
&\leq \frac{\Big(\log(T-2)+1\Big)\Big(20 + \log m + \log n\Big)|A|_{\max}}{T-2}
\end{aligned}$$

which finishes the proof. □

## C. Robust Optimistic Mirror Descent

In this appendix, we prove **Theorem 3**, repeated below for convenience.

**Theorem 2** ($O(\sqrt{T})$-Adversarial Regret)**.** *Suppose that $\|\nabla f_t\|_* \leq G$ for all $t$. Then playing $T$ rounds of **Algorithm 3** with $\eta_t = \frac{1}{G\sqrt{t}}$ against an arbitrary sequence of convex functions has the following guarantee on the regret:*

$$\begin{aligned}
\max_{\mathbf{x}\in\Delta_m}\sum_{t=1}^{T}\langle \mathbf{x}_t - \mathbf{x}, \nabla f_t(\mathbf{x}_t)\rangle &\leq G\sqrt{T}\left(18 + 2D^2\right) + GD\left(3\sqrt{2} + 4D\right) \\
&= O\left(\sqrt{T}\right).
\end{aligned}$$

*Proof.* Define $\mathrm{R}_T^{\mathbf{x}} \coloneqq \sum_{t=1}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, \nabla f_t(\mathbf{x}_t)\rangle$ where $\mathbf{x}^* \coloneqq \arg\min_{\mathbf{x}\in\Delta_m}\Big\langle \mathbf{x}, \sum_{t=1}^{T}\nabla f_t(\mathbf{x}_t)\Big\rangle$. Let $\tilde{\nabla}_t = 2\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})$, and let $\eta_t = \frac{1}{\alpha\sqrt{t}}$ for some $\alpha > 0$ to be chosen later. Then

$$\begin{aligned}
\mathrm{R}_T^{\mathbf{x}} &= \sum_{t=1}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, \nabla f_t(\mathbf{x}_t)\rangle \\
&\leq \sqrt{2}DG + \sum_{t=2}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, \nabla f_t(\mathbf{x}_t) - \tilde{\nabla}_{t-1}\rangle + \sum_{t=2}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, \tilde{\nabla}_{t-1}\rangle \\
&\leq \sqrt{2}DG + \sum_{t=2}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, \nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\rangle - \sum_{t=2}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, \nabla f_{t-1}(\mathbf{x}_{t-1}) - \nabla f_{t-2}(\mathbf{x}_{t-2})\rangle + \sum_{t=2}^{T}\langle \mathbf{x}_t - \mathbf{x}^*, \tilde{\nabla}_{t-1}\rangle \\
&\leq 3\sqrt{2}DG + \sum_{t=2}^{T-1}\langle \mathbf{x}_t - \mathbf{x}_{t+1}, \nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\rangle + \sum_{t=2}^{T}\frac{1}{\eta_t}\Big(D(\mathbf{x}^*, \tilde{\mathbf{x}}_{t-1}) - D(\mathbf{x}^*, \mathbf{x}_t) - D(\mathbf{x}_t, \tilde{\mathbf{x}}_{t-1})\Big) \\
&\leq 3\sqrt{2}DG + \sum_{t=2}^{T-1}\left(\frac{\sqrt{t}G}{9}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2 + \frac{9G}{\sqrt{t}}\right) \\
&\qquad + \alpha\sum_{t=1}^{T}\sqrt{t}\Big(D(\mathbf{x}^*, \tilde{\mathbf{x}}_{t-1}) - D(\mathbf{x}^*, \mathbf{x}_t) - D(\mathbf{x}_t, \tilde{\mathbf{x}}_{t-1})\Big).
\end{aligned}$$

Using the joint convexity of $D(\mathbf{x}, \mathbf{y})$ in $\mathbf{x}$ and $\mathbf{y}$ and the strong convexity of the entropic mirror map, we get:

$$\begin{aligned}
-D(\mathbf{x}_t, \tilde{\mathbf{x}}_{t-1}) &\leq -\frac{1}{2}\|\tilde{\mathbf{x}}_t - \mathbf{x}_{t+1}\|^2 \\
&\leq -\frac{1}{4}\left\|\frac{t-1}{t}(\mathbf{x}_t - \mathbf{x}_{t+1})\right\|^2 + \frac{1}{2}\left(\frac{1}{t}\right)^2\|\mathbf{x}_c - \mathbf{x}_{t+1}\|^2
\end{aligned}$$

$$\leq -\frac{(t-1)^2}{4t^2}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2 + \frac{D^2}{t^2},$$

and

$$D(\mathbf{x}^*, \tilde{\mathbf{x}}_t) \leq \frac{t-1}{t} D(\mathbf{x}^*, \mathbf{x}_t) + \frac{1}{t} D\left(\mathbf{x}^*, \mathbf{x}_c\right).$$

Meanwhile, straightforward calculations show that

$$\sum_{t=2}^{T} \frac{D\left(\mathbf{x}^*, \mathbf{x}_c\right)}{\sqrt{t}} \leq 2D^2\sqrt{T},$$

and

$$\begin{aligned}\sum_{t=2}^{T} \left(\sqrt{t} \cdot \frac{t-1}{t} D(\mathbf{x}^*, \mathbf{x}_{t-1}) - \sqrt{t} D(\mathbf{x}^*, \mathbf{x}_t)\right) &\leq \sum_{t=2}^{T} \left(\sqrt{t-1} D(\mathbf{x}^*, \mathbf{x}_{t-1}) - \sqrt{t} D(\mathbf{x}^*, \mathbf{x}_t)\right) \\ &\leq D(\mathbf{x}^*, \mathbf{x}_1) \leq D^2.\end{aligned}$$

We can hence continue as

$$\begin{aligned}\mathrm{R}_T^{\mathbf{x}} \leq 3\sqrt{2}DG + \sum_{t=2}^{T-1} \left(\frac{\sqrt{t}}{9} G \|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2 + \frac{9G}{\sqrt{t}}\right) + 2\alpha D^2 \sqrt{T} \\ + \alpha D^2 - \frac{\alpha}{4} \sum_{t=2}^{T} \sqrt{t} \cdot \left(\frac{t-1}{t}\right)^2 \|\mathbf{x}_{t-1} - \mathbf{x}_t\|^2 + \alpha D^2 \sum_{t=2}^{T} \frac{\sqrt{t}}{t^2}.\end{aligned} \tag{C.1}$$

Elementary calculations further show

$$\sum_{t=2}^{T-1} \frac{9G}{\sqrt{t}} \leq 18G\sqrt{T},$$

$$\sum_{t=2}^{T} \frac{1}{t\sqrt{t}} \leq 3.$$

Finally, since $(\frac{t-1}{t})^2 \geq \frac{4}{9}$ for $t \geq 3$, we can further bound (C.1) as

$$\begin{aligned}\mathrm{R}_t^{\mathbf{x}} \leq 3\sqrt{2}DG + 18G\sqrt{T} + 2\alpha D^2\sqrt{T} + 4\alpha D^2 \\ + \left(\frac{G}{9} \sum_{t=2}^{T-1} \sqrt{t} \|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2 - \frac{\alpha}{4} \cdot \frac{4}{9} \sum_{t=2}^{T-1} \sqrt{t+1} \|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2\right).\end{aligned}$$

The proof is finished by choosing $\alpha = G$. □

## References

Chiang, Chao-Kai, Yang, Tianbao, Lee, Chia-Jung, Mahdavi, Mehrdad, Lu, Chi-Jen, Jin, Rong, and Zhu, Shenghuo. Online optimization with gradual variations. In *Conference on Learning Theory*, pp. 6–1, 2012.

Rakhlin, Alexander and Sridharan, Karthik. Online learning with predictable sequences. In *Conference on Learning Theory*, pp. 993–1019, 2013.