Last-iterate Convergence in Extensive-Form Games

Chung-Wei Lee
University of Southern California
leechung@usc.edu

Christian Kroer
Columbia University
christian.kroer@columbia.edu

Haipeng Luo
University of Southern California
haipengl@usc.edu

Abstract

Regret-based algorithms are highly efficient at finding approximate Nash equilibria in sequential games such as poker games. However, most regret-based algorithms, including counterfactual regret minimization (CFR) and its variants, rely on iterate averaging to achieve convergence. Inspired by recent advances on *last-iterate* convergence of optimistic algorithms in zero-sum normal-form games, we study this phenomenon in sequential games, and provide a comprehensive study of last-iterate convergence for zero-sum extensive-form games with perfect recall (EFGs), using various optimistic regret-minimization algorithms over treeplexes. This includes algorithms using the vanilla entropy or squared Euclidean norm regularizers, as well as their dilated versions which admit more efficient implementation. In contrast to CFR, we show that all of these algorithms enjoy last-iterate convergence, with some of them even converging *exponentially* fast. We also provide experiments to further support our theoretical results.

1 Introduction

Extensive-form games (EFGs) are an important class of games in game theory and artificial intelligence which can model imperfect information and sequential interactions. EFGs are typically solved by finding or approximating a *Nash equilibrium*. Regret-minimization algorithms are among the most popular approaches to approximate Nash equilibria. The motivation comes from a classical result which says that in two-player zero-sum games, when both players use *no-regret* algorithms, *the average strategy* converges to Nash equilibrium [Freund and Schapire, 1999, Hart and Mas-Colell, 2000, Zinkevich et al., 2007]. Counterfactual Regret Minimization (CFR) [Zinkevich et al., 2007] and it variants such as CFR+ [Tammelin, 2014] are based on this motivation.

However, due to their ergodic convergence guarantee, theoretical convergence rates of regret-minimization algorithms are typically limited to $\mathcal{O}(1/\sqrt{T})$ or $\mathcal{O}(1/T)$ for T rounds, and this is also the case in practice [Brown and Sandholm, 2019a, Burch et al., 2019]. In contrast, it is known that linear convergence rates are achievable for certain other first-order algorithms [Tseng, 1995, Gilpin et al., 2008]. Additionally, the averaging procedure can create complications. It not only increases the computational and memory overhead [Bowling et al., 2015], but also makes things difficult when incorporating neural networks in the solution process, where averaging is usually not possible. Indeed, to address this issue, Brown et al. [2019] create a separate neural network to approximate the average strategy in their Deep CFR model.

Therefore, a natural idea is to design regret-minimization algorithms whose last strategy converges (we call this *last-iterate convergence*), ideally at a faster rate than the average iterate. Unfortunately, many regret minimization algorithms such as regret matching, regret matching+, and hedge, are known not

35th Conference on Neural Information Processing Systems (NeurIPS 2021).

to satisfy this property empirically and theoretically even for normal-form games. Although Bowling et al. [2015] find that in Heads-Up Limit Hold'em poker the last strategy of CFR+ is better than the average strategy, and Farina et al. [2019b] observe in some experiments the last-iterate of optimistic OMD and FTRL converge fast, a theoretical understanding of this phenomenon is still absent for EFGs.

In this work, inspired by recent results on last-iterate convergence in normal-form games [Wei et al., 2021], we greatly extend the theoretical understanding of last-iterate convergence of regret-minimization algorithms in two-player zero-sum extensive-form games with perfect recall, and open up many interesting directions both in theory and practice. First, we show that any optimistic online mirror-descent algorithm instantiated with a strongly convex regularizer that is continuously differentiable on the EFG strategy space provably enjoys last-iterate convergence, while CFR with either regret matching or regret matching+ fails to converge. Moreover, for some of the optimistic algorithms, we further show explicit convergence rates. In particular, we prove that optimistic mirror descent instantiated with the 1-strongly-convex dilated entropy regularizer [Kroer et al., 2020], which we refer to as *Dilated Optimistic Multiplicative Weights Update* (DOMWU), has a *linear* convergence rate under the assumption that there is a unique Nash equilibrium; we note that this assumption was also made by Daskalakis and Panageas [2019], Wei et al. [2021] in order to achieve similar results for normal-form games.

2 Related Work

Extensive-form Games Here we focus on work related to two-player zero-sum perfect-recall games. Although there are many game-solving techniques such as abstraction [Kroer and Sandholm, 2014, Ganzfried and Sandholm, 2014, Brown et al., 2015], endgame solving [Burch et al., 2014, Ganzfried and Sandholm, 2015], and subgame solving [Moravcik et al., 2016, Brown and Sandholm, 2019b], these methods all rely on scalable methods for computing approximate Nash equilibria. There are several classes of algorithms for computing approximate Nash equilibria, such as double-oracle methods [McMahan et al., 2003], fictitious play [Brown, 1951, Heinrich and Silver, 2016], first-order methods [Hoda et al., 2010, Kroer et al., 2020], and CFR methods [Zinkevich et al., 2007, Lanctot et al., 2009, Tammelin, 2014]. Notably, variants of the CFR approach have achieved significant success in poker games [Bowling et al., 2015, Moravčík et al., 2017, Brown and Sandholm, 2018]. Underlying the first-order and CFR approaches is the sequence-form representation [von Stengel, 1996], which allows the problem to be represented as a bilinear saddle-point problem. This leads to algorithms based on smoothing techniques and other first-order methods [Gilpin et al., 2008, Kroer et al., 2017, Gao et al., 2021], and enables CFR via the theorem connecting no-regret guarantees to Nash equilibrium.

Online Convex Optimization and Optimistic Regret Minimization Online convex optimization [Zinkevich, 2003] is a framework for repeated decision making where the goal is to minimize regret. When applied to repeated two-player zero-sum games, it is known that the average strategy converges to Nash Equilibria at the rate of $\mathcal{O}(1/\sqrt{T})$ when both players apply regret-minimization algorithms whose regret grows on the order of $\mathcal{O}(\sqrt{T})$ [Freund and Schapire, 1999, Hart and Mas-Colell, 2000, Zinkevich et al., 2007]. Moreover, when the players use optimistic regret-minimization algorithms, the convergence rate is improved to $\mathcal{O}(1/T)$ [Rakhlin and Sridharan, 2013, Syrgkanis et al., 2015]. Recent works have applied optimism ideas to EFGs, such as optimistic algorithms with dilated regularizers [Kroer et al., 2020, Farina et al., 2019b], CFR-like local optimistic algorithms [Farina et al., 2019a], and optimistic CFR algorithms [Burch, 2018, Brown and Sandholm, 2019a, Farina et al., 2021a]. However, the theoretical results in all these existing papers consider the average strategy, while we are the first to consider last-iterate convergence in EFGs.

Last-iterate Convergence in Saddle-point Optimization As mentioned previously, two-player zero-sum games can be formulated as saddle-point optimization problems. Saddle-point problems have recently gained a lot of attention due to their applications in machine learning, for example in generative adversarial networks [Goodfellow et al., 2014]. Basic algorithms, including gradient descent ascent and multiplicative weights update, diverge even in simple instances [Mertikopoulos et al., 2018, Bailey and Piliouras, 2018]. In contrast, their optimistic versions, optimistic gradient descent ascent (OGDA) [Daskalakis et al., 2018, Mertikopoulos et al., 2019, Wei et al., 2021] and optimistic multiplicative weights update (OMWU) [Daskalakis and Panageas, 2019, Lei et al., 2021, Wei et al., 2021] have been shown to enjoy attractive last-iterate convergence guarantees. However, almost none of these results apply to the case of EFGs: Wei et al. [2021] show a result that implies linear convergence of vanilla OGDA in EFGs (see Corollary 5), but no results are known for vanilla OMWU or more importantly for algorithms instantiated with *dilated* regularizers which lead to fast iterate updates in EFGs. In this work we extend the existing results on normal-form games to EFGs, including the practically-important dilated regularizers.

3 Problem Setup

We start with some basic notation. For a vector $\boldsymbol{z}$, we use z_i to denote its i-th coordinate and $\|\boldsymbol{z}\|_p$ to denote its p-norm (with $\|\boldsymbol{z}\|$ being a shorthand for $\|\boldsymbol{z}\|_2$). For a convex function ψ, the associated Bregman divergence is define as $D_\psi(\boldsymbol{u}, \boldsymbol{v}) = \psi(\boldsymbol{u}) - \psi(\boldsymbol{v}) - \langle \nabla\psi(\boldsymbol{v}), \boldsymbol{u} - \boldsymbol{v}\rangle$, and ψ is called κ-strongly convex with respect to the p-norm if $D_\psi(\boldsymbol{u}, \boldsymbol{v}) \geq \frac{\kappa}{2}\|\boldsymbol{u} - \boldsymbol{v}\|_p^2$ holds for all $\boldsymbol{u}$ and $\boldsymbol{v}$ in the domain. The Kullback-Leibler divergence, which is the Bregman divergence with respect to the entropy function, is denoted by $\mathrm{KL}(\cdot, \cdot)$. Finally, we use Δ_P to denote the $(P-1)$-dimensional simplex and $[N]$ to denote the set $\{1, 2 \ldots, N\}$ for some positive integer N.

Extensive-form Games as Bilinear Saddle-point Optimization We consider the problem of finding a Nash equilibrium of a two-player zero-sum extensive-form game (EFG) with perfect recall. Instead of formally introducing the definition of an EFG (see Appendix A for an example), for the purpose of this work, it suffices to consider an equivalent formulation, which casts the problem as a simple bilinear saddle-point optimization [von Stengel, 1996]:

$$\min_{\boldsymbol{x}\in\mathcal{X}} \max_{y\in\mathcal{Y}} \boldsymbol{x}^\top \boldsymbol{G}\boldsymbol{y} = \max_{y\in\mathcal{Y}} \min_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{x}^\top \boldsymbol{G}\boldsymbol{y}, \tag{1}$$

where $\boldsymbol{G} \in [-1, +1]^{M\times N}$ is a known matrix, and $\mathcal{X} \subset \mathbb{R}^M$ and $\mathcal{Y} \subset \mathbb{R}^N$ are two polytopes called *treeplexes* (to be defined soon). The set of Nash equilibria is then defined as $\mathcal{Z}^* = \mathcal{X}^* \times \mathcal{Y}^*$, where $\mathcal{X}^* = \operatorname{argmin}_{\boldsymbol{x}\in\mathcal{X}} \max_{y\in\mathcal{Y}} \boldsymbol{x}^\top \boldsymbol{G}\boldsymbol{y}$ and $\mathcal{Y}^* = \operatorname{argmax}_{y\in\mathcal{Y}} \min_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{x}^\top \boldsymbol{G}\boldsymbol{y}$. Our goal is to find a point $z \in \mathcal{Z} = \mathcal{X} \times \mathcal{Y}$ that is close to the set of Nash equilibria $\mathcal{Z}^*$, and we use the Bregman divergence (of some function ψ) between $\boldsymbol{z}$ and the closest point in $\mathcal{Z}^*$ to measure the closeness, that is, $\min_{\boldsymbol{z}^*\in\mathcal{Z}^*} D_\psi(\boldsymbol{z}^*, \boldsymbol{z})$.

For notational convenience, we let $P = M + N$ and $F(\boldsymbol{z}) = (\boldsymbol{G}\boldsymbol{y}, -\boldsymbol{G}^\top\boldsymbol{x})$ for any $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{Z} \subset \mathbb{R}^P$. Without loss of generality, we assume $\|F(\boldsymbol{z})\|_\infty \leq 1$ for all $\boldsymbol{z} \in \mathcal{Z}$ (which can always be ensured by normalizing the entries of $\boldsymbol{G}$ accordingly).

Treeplexes The structure of the EFG is implicitly captured by the treeplexes $\mathcal{X}$ and $\mathcal{Y}$, which are generalizations of simplexes that capture the sequential structure of an EFG. The formal definition is as follows. (In Appendix A, we provide more details on the connection between treeplexes and the structure of the EFG, as well as concrete examples of treeplexes for better illustrations.)

Definition 1 (Hoda et al. [2010]). *A treeplex is recursively constructed via the following three operations:*

1. *Every probability simplex is a treeplex.*
2. *Given treeplexes $\mathcal{Z}_1, \ldots, \mathcal{Z}_K$, the Cartesian product $\mathcal{Z}_1 \times \cdots \times \mathcal{Z}_K$ is a treeplex.*
3. *(Branching) Given treeplexes $\mathcal{Z}_1 \subset \mathbb{R}^M$ and $\mathcal{Z}_2 \subset \mathbb{R}^N$, and any $i \in [M]$,*

$$\mathcal{Z}_1 \boxed{i} \mathcal{Z}_2 = \left\{(\boldsymbol{u}, u_i \cdot \boldsymbol{v}) \in \mathbb{R}^{M+N} : \boldsymbol{u} \in \mathcal{Z}_1,\ \boldsymbol{v} \in \mathcal{Z}_2\right\}$$

is a treeplex.

By definition, a treeplex is a tree-like structure built with simplexes, which intuitively represents the tree-like decision space of a single player, and an element in the treeplex represents a strategy for the player. Let $\mathcal{H}^{\mathcal{Z}}$ denote the collection of all the simplexes in treeplex $\mathcal{Z}$, which following Definition 1 can be recursively defined as: $\mathcal{H}^{\mathcal{Z}} = \{\mathcal{Z}\}$ if $\mathcal{Z}$ is a simplex; $\mathcal{H}^{\mathcal{Z}} = \bigcup_{k=1}^K \mathcal{H}^{\mathcal{Z}_i}$ if $\mathcal{Z}$ is a Cartesian product $\mathcal{Z}_1 \times \cdots \times \mathcal{Z}_K$; and $\mathcal{H}^{\mathcal{Z}} = \mathcal{H}^{\mathcal{Z}_1} \cup \mathcal{H}^{\mathcal{Z}_2}$ if $\mathcal{Z} = \mathcal{Z}_1 \boxed{i} \mathcal{Z}_2$. In EFG terminology, $\mathcal{H}^{\mathcal{X}}$ and $\mathcal{H}^{\mathcal{Y}}$ are the collections of *information sets* for player $\boldsymbol{x}$ and player $\boldsymbol{y}$ respectively, which are the decision points for the players, at which they select an action within the simplex. For any $h \in \mathcal{H}^{\mathcal{Z}}$, we let Ω_h denote the set of indices belonging to h, and for any $\boldsymbol{z} \in \mathcal{Z}$, we let $\boldsymbol{z}_h$ be the slice of $\boldsymbol{z}$ whose indices are in Ω_h. For each index i, we also let $h(i)$ be the simplex i falls into, that is, $i \in \Omega_{h(i)}$.

In Definition 1, the last branching operation naturally introduces the concept of a *parent variable* for each $h \in \mathcal{H}^{\mathcal{Z}}$, which can again be recursively defined as: if $\mathcal{Z}$ is a simplex, then it has no parent; if $\mathcal{Z}$ is a Cartesian product $\mathcal{Z}_1 \times \cdots \times \mathcal{Z}_K$, then the parent of $h \in \mathcal{H}^{\mathcal{Z}}$ is the same as the parent of h in the treeplex $\mathcal{Z}_k$ that h belongs to (that is, $h \in \mathcal{H}^{\mathcal{Z}_k}$); finally, if $\mathcal{Z} = \{(\boldsymbol{u}, u_i \cdot \boldsymbol{v}) : \boldsymbol{u} \in \mathcal{Z}_1,\ \boldsymbol{v} \in \mathcal{Z}_2\}$, then for all $h \in \mathcal{H}^{\mathcal{Z}_2}$ without a parent, their parent in $\mathcal{Z}$ is u_i, and for all other h, their parents remain the same as in $\mathcal{Z}_1$ or $\mathcal{Z}_2$. We denote by $\sigma(h)$ the index of the parent variable of h, and let it be 0 if h has no parent. For convenience, we let $z_0 = 1$ for all $z \in \mathcal{Z}$ (so that $z_{\sigma(h)}$ is always well-defined). Also define $\mathcal{H}_i = \{h \in \mathcal{H}^{\mathcal{Z}} : \sigma(h) = i\}$ to be the collection of simplexes whose parent index is i.

Similarly, for an index i, its parent index is defined as $p_i = \sigma(h(i))$, and i is called a *terminal index* if it is not a parent index (that is, $i \neq p_j$ for all j). Finally, for an element $\boldsymbol{z} \in \mathcal{Z}$ and an index i, we define $q_i = z_i / z_{p_i}$. In EFG terminology, q_i specifies the probability of selecting action i in the information set $h(i)$ according to strategy $\boldsymbol{z}$.

4 Optimistic Regret-minimization Algorithms

There are many different algorithms for solving bilinear saddle-point problems over general constrained sets. We focus specifically on a family of regret-minimization algorithms, called Optimistic Online Mirror Descent (OOMD) [Rakhlin and Sridharan, 2013], which are known to be highly efficient. In contrast to the CFR algorithm and its variants, which minimize a local regret notion at each information set (which upper bounds global regret), the algorithms we consider explicitly minimize global regret. As our main results in the next section show, these global regret-minimization algorithms enjoy last-iterate convergence, while CFR provably diverges.

Specifically, given a step size $\eta > 0$ and a convex function ψ (called a *regularizer*), OOMD sequentially performs the following update for $t = 1, 2, \dots,$

$$\boldsymbol{x}_t = \operatorname*{argmin}_{\boldsymbol{x} \in \mathcal{X}} \left\{ \eta \langle \boldsymbol{x}, \boldsymbol{G}\boldsymbol{y}_{t-1} \rangle + D_\psi(\boldsymbol{x}, \widehat{\boldsymbol{x}}_t) \right\}, \qquad \widehat{\boldsymbol{x}}_{t+1} = \operatorname*{argmin}_{\boldsymbol{x} \in \mathcal{X}} \left\{ \eta \langle \boldsymbol{x}, \boldsymbol{G}\boldsymbol{y}_t \rangle + D_\psi(\boldsymbol{x}, \widehat{\boldsymbol{x}}_t) \right\},$$

$$\boldsymbol{y}_t = \operatorname*{argmin}_{\boldsymbol{y} \in \mathcal{Y}} \left\{ \eta \langle \boldsymbol{y}, -\boldsymbol{G}^\top \boldsymbol{x}_{t-1} \rangle + D_\psi(\boldsymbol{y}, \widehat{\boldsymbol{y}}_t) \right\}, \quad \widehat{\boldsymbol{y}}_{t+1} = \operatorname*{argmin}_{\boldsymbol{y} \in \mathcal{Y}} \left\{ \eta \langle \boldsymbol{y}, -\boldsymbol{G}^\top \boldsymbol{x}_t \rangle + D_\psi(\boldsymbol{y}, \widehat{\boldsymbol{y}}_t) \right\},$$

with $(\widehat{\boldsymbol{x}}_1, \widehat{\boldsymbol{y}}_1) = (\boldsymbol{x}_0, \boldsymbol{y}_0) \in \mathcal{Z}$ being arbitrary. Using shorthands $\boldsymbol{z}_t = (\boldsymbol{x}_t, \boldsymbol{y}_t)$, $\widehat{\boldsymbol{z}}_t = (\widehat{\boldsymbol{x}}_t, \widehat{\boldsymbol{y}}_t)$, $\psi(\boldsymbol{z}) = \psi(\boldsymbol{x}) + \psi(\boldsymbol{y})$ and recalling the notation $F(\boldsymbol{z}) = (\boldsymbol{G}\boldsymbol{y}, -\boldsymbol{G}^\top \boldsymbol{x})$, the updates above can be compactly written as OOMD with regularizer ψ over treeplex $\mathcal{Z}$:

$$\boldsymbol{z}_t = \operatorname*{argmin}_{\boldsymbol{z} \in \mathcal{Z}} \left\{ \eta \langle \boldsymbol{z}, F(\boldsymbol{z}_{t-1}) \rangle + D_\psi(\boldsymbol{z}, \widehat{\boldsymbol{z}}_t) \right\}, \ \ \widehat{\boldsymbol{z}}_{t+1} = \operatorname*{argmin}_{\boldsymbol{z} \in \mathcal{Z}} \left\{ \eta \langle \boldsymbol{z}, F(\boldsymbol{z}_t) \rangle + D_\psi(\boldsymbol{z}, \widehat{\boldsymbol{z}}_t) \right\}. \quad (2)$$

Below, we discuss four different regularizers and their resulting algorithms (throughout, we use notations Φ for regularizers based on Euclidean norm and Ψ for regularizers based on entropy).

Vanilla Optimistic Gradient Descent Ascent (VOGDA) Define the vanilla squared Euclidean norm regularizer as $\Phi^{\text{van}}(\boldsymbol{z}) = \frac{1}{2}\sum_i z_i^2$. We call OOMD instantiated with $\psi = \Phi^{\text{van}}$ Vanilla Optimistic Gradient Descent Ascent (VOGDA). In this case, the Bregman divergence is $D_{\Phi^{\text{van}}}(\boldsymbol{z}, \boldsymbol{z}') = \frac{1}{2}\|\boldsymbol{z} - \boldsymbol{z}'\|^2$ (by definition Φ^{van} is thus 1-strongly convex with respect to the 2-norm), and the updates simply become projected gradient descent. For VOGDA there is no closed-form for Eq. (2), since projection onto the treeplex $\mathcal{Z}$ is required. Nevertheless, the solution can still be computed in $\mathcal{O}(P^2 \log P)$ time (recall that P is the dimension of $\mathcal{Z}$) [Gilpin et al., 2008].

Vanilla Optimistic Multiplicative Weight Update (VOMWU) Define the vanilla entropy regularizer as $\Psi^{\text{van}}(\boldsymbol{z}) = \sum_i z_i \ln z_i$. We call OOMD with $\psi = \Psi^{\text{van}}$ Vanilla Optimistic Multiplicative Weights Update (VOMWU). The Bregman divergence in this case is the generalized KL divergence: $D_{\Psi^{\text{van}}}(\boldsymbol{z}, \boldsymbol{z}') = \sum_i z_i \ln(z_i / z_i') - z_i + z_i'$. Although it is well-known that Ψ^{van} is 1-strongly convex with respect to the 1-norm for the special case when $\mathcal{Z}$ is a simplex, this is not true generally on a treeplex. Nevertheless, it can still be shown that Ψ^{van} is 1-strongly convex with respect to the 2-norm; see Appendix C.

The name "Multiplicative Weights Update" is inherited from case when $\mathcal{X}$ and $\mathcal{Y}$ are simplexes, in which case the updates in Eq. (2) have a simple multiplicative form. We emphasize, however, that in general VOMWU does not admit a closed-form update. Instead, to solve Eq. (2), one can equivalently solve a simpler dual optimization problem; see [Zimin and Neu, 2013, Proposition 1].

The two regularizers mentioned above ignore the structure of the treeplex. *Dilated Regularizers* [Hoda et al., 2010], on the other hand, take the structure into account and allow one to decompose the update into simpler updates at each information set. Specifically, given any convex function ψ defined over the simplex and a weight parameter $\boldsymbol{\alpha} \in \mathbb{R}_+^{\mathcal{H}^{\mathcal{Z}}}$, the dilated version of ψ defined over $\mathcal{Z}$ is:

$$\psi_{\boldsymbol{\alpha}}^{\text{dil}}(\boldsymbol{z}) = \sum_{h \in \mathcal{H}^{\mathcal{Z}}} \alpha_h \cdot z_{\sigma(h)} \cdot \psi\left(\frac{\boldsymbol{z}_h}{z_{\sigma(h)}}\right). \tag{3}$$

This is well-defined since ${\boldsymbol{z}_h}/{z_{\sigma(h)}}$ is indeed a distribution within the simplex h (with q_i for $i \in \Omega_h$ being its entries). It can also be shown that $\psi_{\boldsymbol{\alpha}}^{\text{dil}}$ is always convex in $\boldsymbol{z}$ [Hoda et al., 2010]. Intuitively, $\psi_{\boldsymbol{\alpha}}^{\text{dil}}$ applies the base regularizer ψ to the action distribution in each information set and then scales the value by its parent variable and the weight α_h. By picking different base regularizers, we obtain the following two algorithms.

Dilated Optimistic Gradient Descent Ascent (DOGDA) [Farina et al., 2019b] Define the dilated squared Euclidean norm regularizer $\Phi_{\boldsymbol{\alpha}}^{\text{dil}}$ as Eq. (3) with ψ being the vanilla squared Euclidean norm $\psi(\boldsymbol{z}) = \frac{1}{2}\sum_i z_i^2$. Direct calculation shows $\Phi_{\boldsymbol{\alpha}}^{\text{dil}}(\boldsymbol{z}) = \frac{1}{2}\sum_i \alpha_{h(i)} z_i q_i$. We call OOMD with regularizer $\Phi_{\boldsymbol{\alpha}}^{\text{dil}}$ the Dilated Optimistic Gradient Descent Ascent algorithm (DOGDA). It is known that there exists an $\boldsymbol{\alpha}$ such that $\Phi_{\boldsymbol{\alpha}}^{\text{dil}}$ is 1-strongly convex with respect to the 2-norm [Farina et al., 2019b]. Importantly, DOGDA decomposes the update Eq. (2) into simpler gradient descent-style updates at each information set, as shown below.

Lemma 1 (Hoda et al. [2010])**.** *If* $\boldsymbol{z}' = \operatorname{argmin}_{\boldsymbol{z} \in \mathcal{Z}} \left\{\eta\langle \boldsymbol{z}, \boldsymbol{f}\rangle + D_{\Phi_{\boldsymbol{\alpha}}^{dil}}(\boldsymbol{z}, \widehat{\boldsymbol{z}})\right\}$*, then for every* $h \in \mathcal{H}^{\mathcal{Z}}$*, the corresponding vector* $\boldsymbol{q}_h' = \frac{\boldsymbol{z}_h'}{z_{\sigma(h)}'}$ *can be computed by:*

$$\boldsymbol{q}_h' = \operatorname*{argmin}_{\boldsymbol{q}_h \in \Delta_{|\Omega_h|}} \left\{\eta\langle \boldsymbol{q}_h, \boldsymbol{L}_h\rangle + \frac{\alpha_h}{2}\|\boldsymbol{q}_h - \widehat{\boldsymbol{q}}_h\|^2\right\}, \tag{4}$$

where $\widehat{\boldsymbol{q}}_h = \frac{\widehat{\boldsymbol{z}}_h}{\widehat{\boldsymbol{z}}_{\sigma(h)}}$*,* $\boldsymbol{L}_h$ *is the slice of* $\boldsymbol{L}$ *whose entries are in* Ω_h*, and* $\boldsymbol{L}$ *is defined through:*

$$L_i = f_i + \sum_{h \in \mathcal{H}_i} \left(\langle \boldsymbol{q}_h', \boldsymbol{L}_h\rangle + \frac{\alpha_h}{2\eta}\|\boldsymbol{q}_h' - \widehat{\boldsymbol{q}}_h\|^2\right).$$

While the definitions of $\boldsymbol{q}_h'$ and $\boldsymbol{L}$ are seemingly recursive, one can verify that they can in fact be computed in a "bottom-up" manner, starting with the terminal indices. Although Eq. (4) still does not admit closed-form solution, it only requires projection onto a simplex, which can be solved efficiently, see e.g. [Condat, 2016]. Finally, with $\boldsymbol{q}_h'$ computed for all h, $\boldsymbol{z}'$ can be calculated in a "top-down" manner by definition.

Dilated Optimistic Multiplicative Weight Update (DOMWU) [Kroer et al., 2020] Finally, define the dilated entropy regularizer $\Psi_{\boldsymbol{\alpha}}^{\text{dil}}$ as Eq. (3) with ψ being the vanilla entropy $\psi(\boldsymbol{z}) = \sum_i z_i \ln z_i$. Direct calculation shows $\Psi_{\boldsymbol{\alpha}}^{\text{dil}}(\boldsymbol{z}) = \sum_i \alpha_{h(i)} z_i \ln q_i$. We call OOMD with regularizer $\Psi_{\boldsymbol{\alpha}}^{\text{dil}}$ the Dilated Optimistic Multiplicative Weights Update algorithm (DOMWU). Similar to DOGDA, there exists an $\boldsymbol{\alpha}$ such that $\Psi_{\boldsymbol{\alpha}}^{\text{dil}}$ is 1-strongly convex with respect to the 2-norm [Kroer et al., 2020].[1] Moreover, in contrast to all the three algorithms mentioned above, the update of DOMWU has a closed-form solution:

Lemma 2 (Hoda et al. [2010])**.** *Suppose* $\boldsymbol{z}' = \operatorname{argmin}_{\boldsymbol{z} \in \mathcal{Z}} \left\{\eta\langle \boldsymbol{z}, \boldsymbol{f}\rangle + D_{\Psi_{\boldsymbol{\alpha}}^{dil}}(\boldsymbol{z}, \widehat{\boldsymbol{z}})\right\}$*. Similarly to the notation* q_i*, define* $q_i' = z_i'/z_{p_i}'$ *and* $\widehat{q}_i = \widehat{z}_i/\widehat{z}_{p_i}$*. Then we have*

$$q_i' \propto \widehat{q}_i \exp\left(-\eta L_i/\alpha_{h(i)}\right), \text{ where } L_i = f_i - \sum_{h \in \mathcal{H}_i} \frac{\alpha_h}{\eta} \ln\left(\sum_{j \in \Omega_h} \widehat{q}_j \exp\left(-\eta L_j/\alpha_h\right)\right).$$

[1] Kroer et al. [2020] also show a better strong convexity result with respect to the 1-norm. We focus on the 2-norm here for consistency with our other results, but our analysis can be applied to the 1-norm case as well.

This lemma again implies that we can compute q_i' bottom-up, and then $\boldsymbol{z}'$ can be computed top-down. This is similar to DOGDA, except that all updates have a closed-form.

5 Last-iterate Convergence Results

In this section, we present our main last-iterate convergence results for the global regret-minimization algorithms discussed in Section 4. Before doing so, we point out again that the sequence produced by the well-known CFR algorithm may diverge (even if the average converges to a Nash equilibrium). Indeed, this can happen even for a simple normal-form game, as formally shown below.

Theorem 3. *In the rock-paper-scissors game, CFR (with some particular initialization) produces a diverging sequence.*

In fact, we empirically observe that all of CFR, CFR+ [Tammelin, 2014] (with simultaneous updates), and their optimistic versions [Farina et al., 2021a] may diverge in the rock-paper-scissors game. We introduce the algorithms and show the results in Appendix D.

On the contrary, every algorithm from the OOMD family given by Eq. (2) ensures last-iterate convergence, as long as the regularizer is strongly convex and continuously differentiable.

Theorem 4. *Consider the update rules in Eq. (2). Suppose that ψ is 1-strongly convex with respect to the 2-norm and continuously differentiable on the entire domain, and $\eta \le \frac{1}{8P}$. Then $\boldsymbol{z}_t$ converges to a Nash equilibrium as $t \to \infty$.*

As mentioned, Φ^{van} and Ψ^{van} are both 1-strongly convex with respect to 2-norm, so are $\Phi^{\text{dil}}_{\boldsymbol{\alpha}}$ and $\Psi^{\text{dil}}_{\boldsymbol{\alpha}}$ under some specific choice of $\boldsymbol{\alpha}$ (in the rest of the paper, we fix this choice of $\boldsymbol{\alpha}$). However, only Φ^{van} and $\Phi^{\text{dil}}_{\boldsymbol{\alpha}}$ are continuously differentiable in the entire domain. Therefore, Theorem 4 provides an asymptotic convergence result only for VOGDA and DOGDA, but not VOMWU and DOMWU. Nevertheless, below, we resort to different analyses to show a concrete last-iterate convergence rate for three of our algorithms, which is a much more challenging task.

First of all, note that [Wei et al., 2021, Theorem 5, Theorem 8] already provide a general last-iterate convergence rate for VOGDA over polytopes. Since treeplexes are polytopes, we can directly apply their results and obtain the following corollary.

Corollary 5. *Define $\text{dist}^2(\boldsymbol{z}, \mathcal{Z}^*) = \min_{\boldsymbol{z}^* \in \mathcal{Z}^*} \|\boldsymbol{z} - \boldsymbol{z}^*\|^2$. For $\eta \le \frac{1}{8P}$,* VOGDA *guarantees*

$$\text{dist}^2(\boldsymbol{z}_t, \mathcal{Z}^*) \le 64\text{dist}^2(\widehat{\boldsymbol{z}}_1, \mathcal{Z}^*)(1 + C_1)^{-t},$$

where $C_1 > 0$ is some constant that depends on the game and η.

However, the results for VOMWU in [Wei et al., 2021, Theorem 3] is very specific to normal-form game (that is, when $\mathcal{X}$ and $\mathcal{Y}$ are simplexes) and thus cannot be applied here. Nevertheless, we are able to extend their analysis to get the following result.

Theorem 6. *If the EFG has a unique Nash equilibrium $\boldsymbol{z}^*$, then* VOMWU *with step size $\eta \le \frac{1}{8P}$ guarantees $\frac{1}{2}\|\widehat{\boldsymbol{z}}_t - \boldsymbol{z}^*\|^2 \le D_{\Psi^{\text{van}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_t) \le \frac{C_2}{t}$, where $C_2 > 0$ is some constant depending on the game, $\widehat{\boldsymbol{z}}_1$, and η.*

We note that the uniqueness assumption is often required in the analysis of OMWU even for normal-form games [Daskalakis and Panageas, 2019, Wei et al., 2021] (although [Wei et al., 2021, Appendix A.5] provides empirical evidence to show that this may be an artifact of the analysis). Also note that for normal-form games, [Wei et al., 2021, Theorem 3] show a linear convergence rate, whereas here we only show a slower sub-linear rate, due to additional complications introduced by treeplexes (see more discussions in the next section). Whether this can be improved is left as a future direction.

On the other hand, thanks to the closed-form updates of DOMWU, we are able to show the following linear convergence rate for this algorithm.

Theorem 7. *If the EFG has a unique Nash equilibrium $\boldsymbol{z}^*$, then* DOMWU *with step size $\eta \le \frac{1}{8P}$ guarantees $\frac{1}{2}\|\boldsymbol{z}_t - \boldsymbol{z}^*\|^2 \le D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}^*, \boldsymbol{z}_t) \le C_3(1 + C_4)^{-t}$, where $C_3, C_4 > 0$ are constants that depend on the game, $\widehat{\boldsymbol{z}}_1$, and η.*

To the best of our knowledge, this is the first last-iterate convergence result for algorithms with dilated regularizers. Unfortunately, due to technical difficulties, we were unable to prove similar results for DOGDA (see Appendix E for more discussion). We leave that as an important future direction.

6 Analysis Overview

In this section, we provide an overview of our analysis. It starts from the following standard one-step regret analysis of OOMD (see, for example, [Wei et al., 2021, Lemma 1]):

Lemma 8. *Consider the update rules in Eq. (2). Suppose that ψ is 1-strongly convex with respect to the 2-norm, $\|F(\boldsymbol{z}_1) - F(\boldsymbol{z}_2)\| \le L\|\boldsymbol{z}_1 - \boldsymbol{z}_2\|$ for all $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathcal{Z}$ and some $L > 0$, and $\eta \le \frac{1}{8L}$. Then for any $\boldsymbol{z} \in \mathcal{Z}$ and any $t \ge 1$, we have*

$$\eta F(\boldsymbol{z}_t)^\top(\boldsymbol{z}_t - \boldsymbol{z}) \le D_\psi(\boldsymbol{z}, \widehat{\boldsymbol{z}}_t) - D_\psi(\boldsymbol{z}, \widehat{\boldsymbol{z}}_{t+1}) - D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) - \tfrac{15}{16}D_\psi(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t) + \tfrac{1}{16}D_\psi(\widehat{\boldsymbol{z}}_t, \boldsymbol{z}_{t-1}).$$

Note that the Lipschitz condition on F holds in our case with $L = P$ since

$$\|F(\boldsymbol{z}_1) - F(\boldsymbol{z}_2)\| = \sqrt{\|\boldsymbol{G}(\boldsymbol{y}_1 - \boldsymbol{y}_2)\|^2 + \|\boldsymbol{G}^\top(\boldsymbol{x}_1 - \boldsymbol{x}_2)\|^2} \le \sqrt{P\|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_1^2} \le P\|\boldsymbol{z}_1 - \boldsymbol{z}_2\|,$$

which is also why the step size is chosen to be $\eta \le \frac{1}{8P}$ in all our results. In the following, we first prove Theorem 4. Then, we review the convergence analysis of [Wei et al., 2021] for OMWU in normal-form games, and finally demonstrate how to prove Theorem 6 and Theorem 7 by building upon this previous work and addressing the additional complications from EFGs.

6.1 Proof of Theorem 4

For any $\boldsymbol{z}^* \in \mathcal{Z}^*$, by optimality of $\boldsymbol{z}^*$ we have:

$$F(\boldsymbol{z}_t)^\top(\boldsymbol{z}_t - \boldsymbol{z}^*) = \boldsymbol{x}_t^\top\boldsymbol{G}\boldsymbol{y}_t - \boldsymbol{x}_t^\top\boldsymbol{G}\boldsymbol{y}_t + \boldsymbol{x}_t^\top\boldsymbol{G}\boldsymbol{y}^* - {\boldsymbol{x}^*}^\top\boldsymbol{G}\boldsymbol{y}_t \ge {\boldsymbol{x}^*}^\top\boldsymbol{G}\boldsymbol{y}^* - {\boldsymbol{x}^*}^\top\boldsymbol{G}\boldsymbol{y}^* = 0.$$

Thus, taking $\boldsymbol{z} = \boldsymbol{z}^*$ in Lemma 8 and rearranging, we arrive at

$$D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1}) \le D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_t) - D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) - \tfrac{15}{16}D_\psi(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t) + \tfrac{1}{16}D_\psi(\widehat{\boldsymbol{z}}_t, \boldsymbol{z}_{t-1}).$$

Defining $\Theta_t = D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_t) + \frac{1}{16}D_\psi(\widehat{\boldsymbol{z}}_t, \boldsymbol{z}_{t-1})$ and $\zeta_t = D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) + D_\psi(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t)$, we rewrite the inequality above as

$$\Theta_{t+1} \le \Theta_t - \tfrac{15}{16}\zeta_t. \tag{5}$$

We remark that similar inequalities appear in [Wei et al., 2021, Eq. (3) and Eq. (4)], but here we pick $\boldsymbol{z}^* \in \mathcal{Z}^*$ arbitrarily while they have to pick a particular $\boldsymbol{z}^* \in \mathcal{Z}^*$ (such as the projection of $\widehat{\boldsymbol{z}}_t$ onto $\mathcal{Z}^*$). Summing Eq. (5) over t, telescoping, and applying the strong convexity of ψ, we have

$$\Theta_1 \ge \Theta_1 - \Theta_T \ge \frac{15}{16}\sum_{t=1}^{T-1}\zeta_t \ge \frac{15}{32}\sum_{t=1}^{T-1}\|\widehat{\boldsymbol{z}}_{t+1} - \boldsymbol{z}_t\|^2 + \|\boldsymbol{z}_t - \widehat{\boldsymbol{z}}_t\|^2 \ge \frac{15}{64}\sum_{t=2}^{T-1}\|\boldsymbol{z}_t - \boldsymbol{z}_{t-1}\|^2.$$

Similar to the last inequality, we also have $\Theta_1 \ge \frac{15}{64}\sum_{t=1}^{T-1}\|\widehat{\boldsymbol{z}}_{t+1} - \widehat{\boldsymbol{z}}_t\|^2$ since $2\|\widehat{\boldsymbol{z}}_{t+1} - \boldsymbol{z}_t\|^2 + 2\|\boldsymbol{z}_t - \widehat{\boldsymbol{z}}_t\|^2 \ge \|\widehat{\boldsymbol{z}}_{t+1} - \widehat{\boldsymbol{z}}_t\|^2$. Therefore, we conclude that $\|\boldsymbol{z}_t - \widehat{\boldsymbol{z}}_t\|$, $\|\boldsymbol{z}_{t+1} - \boldsymbol{z}_t\|$, and $\|\widehat{\boldsymbol{z}}_{t+1} - \widehat{\boldsymbol{z}}_t\|$ all converge to 0 as $t \to \infty$. On the other hand, since the sequence $\{\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots,\}$ is bounded, by the Bolzano-Weierstrass theorem, there exists a convergent subsequence, which we denote by $\{\boldsymbol{z}_{i_1}, \boldsymbol{z}_{i_2}, \ldots,\}$. Let $\boldsymbol{z}_\infty = \lim_{\tau\to\infty}\boldsymbol{z}_{i_\tau}$. By $\|\widehat{\boldsymbol{z}}_t - \boldsymbol{z}_t\| \to 0$ we also have $\boldsymbol{z}_\infty = \lim_{\tau\to\infty}\widehat{\boldsymbol{z}}_{i_\tau}$. Now, using the first-order optimality condition of $\widehat{\boldsymbol{z}}_{t+1}$, we have for every $\boldsymbol{z}' \in \mathcal{Z}$,

$$(\nabla\psi(\widehat{\boldsymbol{z}}_{t+1}) - \nabla\psi(\widehat{\boldsymbol{z}}_t) + \eta F(\boldsymbol{z}_t))^\top(\boldsymbol{z}' - \widehat{\boldsymbol{z}}_{t+1}) \ge 0.$$

Apply this with $t = i_\tau$ for every τ and let $\tau \to \infty$, we obtain

$$\begin{aligned}
0 &\le \lim_{\tau\to\infty}(\nabla\psi(\widehat{\boldsymbol{z}}_{i_\tau+1}) - \nabla\psi(\widehat{\boldsymbol{z}}_{i_\tau}) + \eta F(\boldsymbol{z}_{i_\tau}))^\top(\boldsymbol{z}' - \widehat{\boldsymbol{z}}_{i_\tau+1}) && \text{(by the first-order optimality)} \\
&= \lim_{\tau\to\infty}\eta F(\boldsymbol{z}_{i_\tau})^\top(\boldsymbol{z}' - \widehat{\boldsymbol{z}}_{i_\tau+1}) && \text{(by } \|\widehat{\boldsymbol{z}}_{t+1} - \widehat{\boldsymbol{z}}_t\| \to 0 \text{ and the continuity of } \nabla\psi) \\
&= \eta F(\boldsymbol{z}_\infty)^\top(\boldsymbol{z}' - \boldsymbol{z}_\infty) && \text{(by } \boldsymbol{z}_\infty = \lim_{\tau\to\infty}\boldsymbol{z}_{i_\tau} = \lim_{\tau\to\infty}\widehat{\boldsymbol{z}}_{i_\tau})
\end{aligned}$$

This implies that $\boldsymbol{z}_\infty$ is a Nash equilibrium. Finally, choosing $\boldsymbol{z}^* = \boldsymbol{z}_\infty$ in the definition of Θ_t, we have $\lim_{\tau\to\infty} \Theta_{i_\tau} = 0$ because $\lim_{\tau\to\infty} D_\psi(\boldsymbol{z}_\infty, \widehat{\boldsymbol{z}}_{i_\tau}) = 0$ and $\lim_{\tau\to\infty} \|\widehat{\boldsymbol{z}}_{i_\tau} - \boldsymbol{z}_{i_\tau - 1}\| = 0$. Additionally, by Eq. (5) we also have that $\lim_{t\to\infty} \Theta_t = 0$ as Θ_t is non-increasing. Therefore, we conclude that the entire sequence $\{\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots\}$ converges to $\boldsymbol{z}_\infty$. On the other hand, since OOMD is a regret-minimization algorithm, it is well known that the average iterate converges to a Nash equilibrium [Freund and Schapire, 1999]. Consequently, combining the two facts above implies that $\boldsymbol{z}_t$ has to converge to a Nash equilibrium, which proves Theorem 4.

We remark that Lemma 8 holds for general closed convex domains as shown in [Wei et al., 2021]. Consequently, with the same argument, Theorem 4 holds more generally as long as $\mathcal{X}$ and $\mathcal{Y}$ are closed convex sets. While the argument is straightforward, we are not aware of similar results in prior works. Also note that unlike Theorem 6 and Theorem 7, Theorem 4 holds without the uniqueness assumption for VOMWU and DOMWU.

6.2 Review for normal-form games

To better explain our analysis and highlight its novelty, we first review the two-stage analysis of [Wei et al., 2021] for OMWU in normal-form games, a special case of our setting when $\mathcal{X}$ and $\mathcal{Y}$ are simplexes. Note that both VOMWU and DOMWU reduce to OMWU in this case. As with Theorem 6 and Theorem 7, the normal-form OMWU results assume a unique Nash equilibrium $\boldsymbol{z}^*$. With this uniqueness assumption and [Mertikopoulos et al., 2018, Lemma C.4], Wei et al. [2021] show the following inequality

$$\zeta_t = D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) + D_\psi(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t) \geq C_5 \|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_{t+1}\|^2 \tag{6}$$

for some problem-dependent constant $C_5 > 0$, which, when combined with Eq. (5), implies that if the algorithm's current iterate is far from $\boldsymbol{z}^*$, then the decrease in Θ_t is more substantial, that is, the algorithm makes more progress on approaching $\boldsymbol{z}^*$. To establish a recursion, however, we need to connect the 2-norm back to the Bregman divergence (a reverse direction of strong convexity). To do so, Wei et al. [2021] argue that $\widehat{z}_{t+1,i}$ can be lower bounded by another problem-dependent constant for $i \in \mathrm{supp}(\boldsymbol{z}^*)$ [Wei et al., 2021, Lemma 19], where $\mathrm{supp}(\boldsymbol{z}^*)$ denotes the support of $\boldsymbol{z}^*$. This further allows them to lower bound $\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_{t+1}\|$ in terms of $D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1})$ (which is just $\mathrm{KL}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1})$), leading to

$$\zeta_t = D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) + D_\psi(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t) \geq C_6 D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1})^2, \tag{7}$$

for some $C_6 > 0$. On the other hand, ignoring the nonnegative term $D_\psi(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t)$, we also have:

$$\zeta_t = D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) + D_\psi(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t) \geq D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) \geq \frac{1}{4} D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t)^2, \tag{8}$$

where the last step uses the fact that $\widehat{\boldsymbol{z}}_{t+1}$ and $\boldsymbol{z}_t$ are close [Wei et al., 2021, Lemma 17 and Lemma 18]. Now, Eq. (7) and Eq. (8) together imply $6\zeta_t \geq 2C_6 D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1})^2 + D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t)^2 \geq \min\left\{C_6, \frac{1}{2}\right\} \Theta_{t+1}^2$. Plugging this back into Eq. (5), we obtain a recursion

$$\Theta_{t+1} \leq \Theta_t - C_7 \Theta_{t+1}^2 \tag{9}$$

for some $C_7 > 0$, which then implies $\Theta_t = O(1/t)$ [Wei et al., 2021, Lemma 12]. This can be seen as the first and slower stage of the convergence behavior of the algorithm.

To further show a linear convergence rate, they argue that there exists a constant $C_8 > 0$ such that when the algorithm's iterate is reasonably close to $\boldsymbol{z}^*$ in the following sense:

$$\max\{\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_t\|_1, \|\boldsymbol{z}^* - \boldsymbol{z}_t\|_1\} \leq C_8, \tag{10}$$

the following improved version of Eq. (7) holds (note the lack of square on the right-hand side):

$$\zeta_t = D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) + D_\psi(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t) \geq C_9 D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1}) \tag{11}$$

for some constant $0 < C_9 < 1$. Therefore, using the $1/t$ convergence rate derived in the first stage, there exists a T_0 such that when $t \geq T_0$, Eq. (10) holds and the algorithm enters the second stage. In this stage, combining Eq. (11) and the fact $\zeta_t \geq D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t)$ gives $\zeta_t \geq \frac{C_9}{2}\Theta_{t+1}$, which, together with Eq. (5) again, implies an improved recursion $\Theta_{t+1} \leq \Theta_t - \frac{15}{32} C_9 \Theta_{t+1}$. This finally shows a linear convergence rate $\Theta_t = O((1+\rho)^{-t})$ for some problem-dependent constant $\rho > 0$.

6.3 Analysis of Theorem 6 and Theorem 7

While we mainly follow the steps of the analysis of [Wei et al., 2021] discussed above to prove Theorem 6 and Theorem 7, we remark that the generalization is highly non-trivial. First of all, we have to prove Eq. (6) for $\mathcal{Z}$ being a general treeplex, which does not follow [Mertikopoulos et al., 2018, Lemma C.4] since its proof is very specific to simplexes. Instead, we prove it by writing down the primal-dual linear program of Eq. (1) and applying the strict complementary slackness; see Appendix E.1 for details.

Next, to connect the 2-norm back to the Bregman divergence (which is not the simple KL divergence anymore, especially for DOMWU), we prove the following for VOMWU:

$$D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1}) \le \sum_{i\in\text{supp}(\boldsymbol{z}^*)} \frac{(z_i^* - \widehat{z}_{t+1,i})^2}{\widehat{z}_{t+1,i}} + \sum_{i\notin\text{supp}(\boldsymbol{z}^*)} \widehat{z}_{t+1,i} \le \frac{3P\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_{t+1}\|}{\min_{i\in\text{supp}(\boldsymbol{z}^*)} \widehat{z}_{t+1,i}}, \tag{12}$$

and the following for DOMWU:

$$\frac{D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1})}{C'} \le \sum_{i\in\text{supp}(\boldsymbol{z}^*)} \frac{(z_i^* - \widehat{z}_{t+1,i})^2}{z_i^* \widehat{q}_{t+1,i}} + \sum_{i\notin\text{supp}(\boldsymbol{z}^*)} z_{p_i}^* \widehat{q}_{t+1,i} \le \frac{\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_{t+1}\|_1}{\min_{i\in\text{supp}(\boldsymbol{z}^*)} z_i^* \widehat{z}_{t+1,i}}, \tag{13}$$

where $C' = 4P\|\boldsymbol{\alpha}\|_\infty$ (see Appendix E.2). We then show a lower bound on $z_{t+1,i}$ and $\widehat{z}_{t+1,i}$ for all $i \in \text{supp}(z^*)$, using similar arguments of [Wei et al., 2021] (see Appendix E.3). Combining Eq. (12) and Eq. (13) with Eq. (6), we have the counterpart of Eq. (7) for both VOMWU and DOMWU.

Showing Eq. (8) also involves extra complication if we follow their analysis, especially for VOMWU which does not admit a closed-form update. Instead, we find a simple workaround: by applying Eq. (5) repeatedly, we get $D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1) = \Theta_1 \ge \cdots \ge \Theta_{t+1} \ge \frac{1}{16} D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t)$, thus, $\zeta_t \ge D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) \ge C_{10} D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t)^2$ for some $C_{10} > 0$ depending on $D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1)$. Combining this with Eq. (7), and applying them to Eq. (5), we obtain the recursion $\Theta_{t+1} \le \Theta_t - C_{11}\Theta_{t+1}^2$ for some $C_{11} > 0$ similar to Eq. (9), which implies $\Theta_t = O(1/t)$ for both VOMWU and DOMWU and proves Theorem 6.

Finally, to show a linear convergence rate, we need to show the counterpart of Eq. (11), which is again more involved compared to the normal-form game case. Indeed, we are only able to do so for DOMWU by making use of its closed-form update described in Lemma 2. Specifically, observe that in Eq. (13), the term $\sum_{i\notin\text{supp}(\boldsymbol{z}^*)} z_{p_i}^* \widehat{q}_{t+1,i}$ is the one that prevents us from bounding $D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1})$ by $\mathcal{O}(\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_{t+1}\|^2)$. Thus, our high-level idea is to argue that $\sum_{i\notin\text{supp}(\boldsymbol{z}^*)} z_{p_i}^* \widehat{q}_{t+1,i}$ decreases significantly as $\widehat{\boldsymbol{z}}_t$ gets close enough to $\boldsymbol{z}^*$. To do so, we use a bottom-up induction to prove that, for any information set $h \in \mathcal{H}^{\mathcal{Z}}$, indices $i, j \in \Omega_h$ such that $i \notin \text{supp}(\boldsymbol{z}^*)$ and $j \in \text{supp}(\boldsymbol{z}^*)$, $\widehat{L}_{t,i}$ is significantly larger than $\widehat{L}_{t,j}$ when $\widehat{\boldsymbol{z}}_t$ is close to $\boldsymbol{z}^*$, where $\widehat{\boldsymbol{L}}_t$ is the counterpart of $\boldsymbol{L}$ in Lemma 2 when computing of $\widehat{\boldsymbol{q}}_{t+1}$. This makes sure that the term $\sum_{i\notin\text{supp}(\boldsymbol{z}^*)} z_{p_i}^* \widehat{q}_{t+1,i}$ is dominated by the other term involving $i \in \text{supp}(\boldsymbol{z}^*)$ in Eq. (13), which eventually helps us show Eq. (11) and the final linear convergence rate in Theorem 7. See Appendix E.5 for details.

7 Experiments

In this section, we experimentally evaluate the algorithms on three standard EFG benchmarks: Kuhn poker [Kuhn, 1950], Pursuit-evasion [Kroer et al., 2018], and Leduc poker [Southey et al., 2005]. The results are shown in Figure 1. Besides the optimistic algorithms, we also show two CFR-based algorithms as reference points. "CFR+" refers to CFR with alternating updates, linear averaging [Tammelin, 2014], and regret matching+ as the regret minimizer. "CFR w/ RM+" is CFR with regret matching+ and linear averaging. We provide the formal descriptions of these two algorithms in Appendix D for completeness. For the optimistic algorithms, we plot the last iterate performance. For the CFR-based algorithms, we plot the performance of the linear average of iterates (recall that the last iterate of CFR-based algorithms is not guaranteed to converge to a Nash equilibrium).

For Kuhn poker and Pursuit-evasion (on the left and in the middle of Figure 1), all of the optimistic algorithms perform much better than CFR+, and their curves are nearly straight, showing their linear last-iterate convergence on these games.

For Leduc poker, although CFR+ performs the best, we can still observe the last-iterate convergence trends of the optimistic algorithms. We remark that although VOGDA and DOMWU have linear

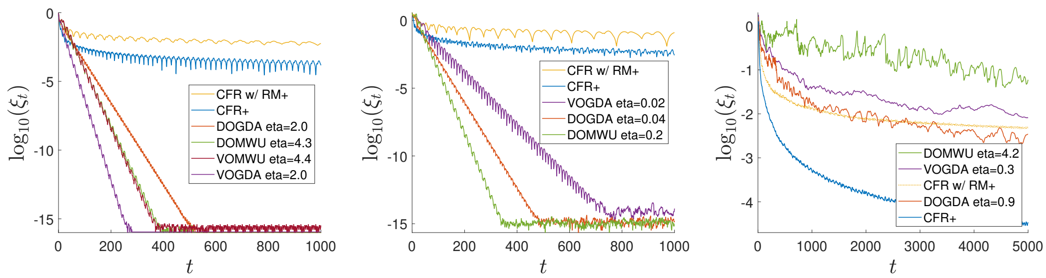

Figure 1: Experiments on Kuhn poker (left), Pursuit-evasion (middle), and Leduc poker (right). A description of each game is given in Appendix B. $\xi_t = \max_{\boldsymbol{y}} \overline{\boldsymbol{x}}_t^\top \boldsymbol{G}\boldsymbol{y} - \min_{\boldsymbol{x}} \boldsymbol{x}^\top \boldsymbol{G}\overline{\boldsymbol{y}}_t$ is the duality gap at time step t, where $(\overline{\boldsymbol{x}}_t, \overline{\boldsymbol{y}}_t)$ is the approximate Nash equilibrium computed by the algorithm at time t (for the optimistic algorithms, $(\overline{\boldsymbol{x}}_t, \overline{\boldsymbol{y}}_t)$ is $(\boldsymbol{x}_t, \boldsymbol{y}_t)$ while for the CFR-based algorithms, $(\overline{\boldsymbol{x}}_t, \overline{\boldsymbol{y}}_t)$ is the linear average). The legend order reflects the curve order at the right-most point. Due to much higher computation overhead than all the other algorithms, we only run VOMWU on Kuhn poker, the game with the smallest size among the three games. For each optimistic algorithm, we fine-tune step size η to get better convergence results and show its value in the legends. There is no hyperparameter for the CFR-based algorithms. All the experiments are run on CPU in a personal computer and the total computation time is less than an hour. There is no random seed and the results are all deterministic.

convergence rate in theory, the experiment on Leduc uses a step size η which is much larger than Corollary 5 and Theorem 7 suggest, which may void the linear convergence guarantee. This is done because the theoretical step size takes too many iterations before it starts to show improvement. It is worth noting that CFR+ improves significantly when changing simultaneous updates (that is, CFR w/ RM+) to alternating updates. Analyzing alternation and combining it with optimistic algorithms is a promising direction. We provide a description of each game, more discussions, and details of the experiments in Appendix B.

8 Conclusions

In this work, we developed the first general last-iterate convergence results for solving EFGs. Our paper opens up many potential future directions. The recent dilatable global entropy regularizer of Farina et al. [2021b] can likely be analyzed using techniques similar to our analysis of VOMWU and DOMWU, and it would likely lead to a linear rate as with DOMWU, due to its closed-form DOMWU-style update. Other natural questions include whether it is possible to obtain better convergence rates for VOMWU and DOGDA, whether one can remove the uniqueness assumption for VOMWU and DOMWU, and finally whether it is possible to obtain last-iterate convergence rates for CFR-like optimistic algorithms such as those in [Farina et al., 2021a]. On the practical side, optimistic algorithms with last-iterate convergence guarantees may allow more efficient computation and better incorporation with deep learning-based game-solving approaches.

Acknowledgments and Disclosure of Funding

CWL and HL are supported by NSF Award IIS-1943607.

References

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

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

George W Brown. Iterative solution of games by fictitious play. *Activity analysis of production and allocation*, 13(1):374–376, 1951.

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. Solving imperfect-information games via discounted regret minimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 1829–1836, 2019a.

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

Noam Brown, Sam Ganzfried, and Tuomas Sandholm. Hierarchical abstraction, distributed equilibrium computation, and post-processing, with application to a champion no-limit texas hold'em agent. In *AAAI Workshop: Computer Poker and Imperfect Information*, 2015.

Noam Brown, Adam Lerer, Sam Gross, and Tuomas Sandholm. Deep counterfactual regret minimization. In *International conference on machine learning*, pages 793–802. PMLR, 2019.

Neil Burch. *Time and space: Why imperfect information games are hard*. PhD thesis, University of Alberta, 2018.

Neil Burch, Michael Johanson, and Michael Bowling. Solving imperfect information games using decomposition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, 2014.

Neil Burch, Matej Moravcik, and Martin Schmid. Revisiting CFR+ and alternating updates. *Journal of Artificial Intelligence Research*, 64:429–443, 2019.

Laurent Condat. Fast projection onto the simplex and the ℓ_1 ball. *Mathematical Programming*, 158 (1):575–585, 2016.

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

Constantinos Daskalakis, Andrew Ilyas, Vasilis Syrgkanis, and Haoyang Zeng. Training gans with optimism. In *International Conference on Learning Representations*, 2018.

Gabriele Farina, Christian Kroer, Noam Brown, and Tuomas Sandholm. Stable-predictive optimistic counterfactual regret minimization. In *International conference on machine learning*, pages 1853–1862. PMLR, 2019a.

Gabriele Farina, Christian Kroer, and Tuomas Sandholm. Optimistic regret minimization for extensive-form games via dilated distance-generating functions. In *Advances in Neural Information Processing Systems*, pages 5222–5232, 2019b.

Gabriele Farina, Christian Kroer, and Tuomas Sandholm. Faster game solving via predictive Blackwell approachability: Connecting regret matching and mirror descent. In *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI, 2021a.

Gabriele Farina, Christian Kroer, and Tuomas Sandholm. Better regularization for sequential decision spaces: Fast convergence rates for nash, correlated, and team equilibria. In *Proceedings of the 2021 ACM Conference on Economics and Computation*, 2021b.

Yoav Freund and Robert E Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1-2):79–103, 1999.

Sam Ganzfried and Tuomas Sandholm. Potential-aware imperfect-recall abstraction with earth mover's distance in imperfect-information games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, 2014.

Sam Ganzfried and Tuomas Sandholm. Endgame solving in large imperfect-information games. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 37–45. Citeseer, 2015.

Yuan Gao, Christian Kroer, and Donald Goldfarb. Increasing iterate averaging for solving saddle-point problems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021.

Andrew Gilpin, Javier Peña, and Tuomas Sandholm. First-order algorithm with $\mathcal{O}(\ln(1/\epsilon))$ convergence for ϵ-equilibrium in two-person zero-sum games. In *AAAI*, 2008.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, 2014.

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

Johannes Heinrich and David Silver. Deep reinforcement learning from self-play in imperfect-information games. *arXiv preprint arXiv:1603.01121*, 2016.

Samid Hoda, Andrew Gilpin, Javier Pena, and Tuomas Sandholm. Smoothing techniques for computing nash equilibria of sequential games. *Mathematics of Operations Research*, 35(2): 494–512, 2010.

Christian Kroer and Tuomas Sandholm. Extensive-form game abstraction with bounds. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 621–638, 2014.

Christian Kroer, Gabriele Farina, and Tuomas Sandholm. Smoothing method for approximate extensive-form perfect equilibrium. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 295–301, 2017.

Christian Kroer, Gabriele Farina, and Tuomas Sandholm. Robust stackelberg equilibria in extensive-form games and extension to limited lookahead. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

Christian Kroer, Kevin Waugh, Fatma Kılınç-Karzan, and Tuomas Sandholm. Faster algorithms for extensive-form game solving via improved smoothing functions. *Mathematical Programming*, 179 (1):385–417, 2020.

HW Kuhn. Simplified two-person poker. *Contributions to the Theory of Games*, pages 97–103, 1950.

Marc Lanctot, Kevin Waugh, Martin Zinkevich, and Michael H Bowling. Monte carlo sampling for regret minimization in extensive games. In *NIPS*, pages 1078–1086, 2009.

Qi Lei, Sai Ganesh Nagarajan, Ioannis Panageas, and Xiao Wang. Last iterate convergence in no-regret learning: constrained min-max optimization for convex-concave landscapes. *The 24nd International Conference on Artificial Intelligence and Statistics*, 2021.

H Brendan McMahan, Geoffrey J Gordon, and Avrim Blum. Planning in the presence of cost functions controlled by an adversary. In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pages 536–543, 2003.

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*, 2018.

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

Matej Moravcik, Martin Schmid, Karel Ha, Milan Hladik, and Stephen Gaukrodger. Refining subgames in large imperfect information games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30, 2016.

Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisỳ, 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.

Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani. *Algorithmic Game Theory*. Cambridge University Press, 2007.

Sasha Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. In *Advances in Neural Information Processing Systems*, pages 3066–3074, 2013.

Finnegan Southey, Michael Bowling, Bryce Larson, Carmelo Piccione, Neil Burch, Darse Billings, and Chris Rayner. Bayes' bluff: opponent modelling in poker. In *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence*, pages 550–558, 2005.

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

Oskari Tammelin. Solving large imperfect information games using CFR+. *arXiv preprint arXiv:1407.5042*, 2014.

Paul Tseng. On linear convergence of iterative methods for the variational inequality problem. In *Journal of Computational and Applied Mathematics*, 1995.

Robert J Vanderbei et al. *Linear programming*, volume 3. Springer, 2015.

Bernhard von Stengel. Efficient computation of behavior strategies. *Games and Economic Behavior*, 14(2):220–246, 1996.

Chen-Yu Wei, Chung-Wei Lee, Mengxiao Zhang, and Haipeng Luo. Linear last-iterate convergence in constrained saddle-point optimization. In *International Conference on Learning Representations*, 2021.

Alexander Zimin and Gergely Neu. Online learning in episodic Markovian decision processes by relative entropy policy search. In *Neural Information Processing Systems 26*, 2013.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th international conference on machine learning (icml-03)*, pages 928–936, 2003.

Martin Zinkevich, Michael Johanson, Michael Bowling, and Carmelo Piccione. Regret minimization in games with incomplete information. *Advances in neural information processing systems*, 20: 1729–1736, 2007.

A Examples of EFG and Treeplexes

In this section, we introduce Kuhn poker [Kuhn, 1950], a simple EFG as an example to introduce treeplexes and the corresponding definitions. In this game, there are three cards in the deck: King, Queen, and Jack. Both player $\boldsymbol{x}$ and player $\boldsymbol{y}$ are dealt one card, while the third card is put aside unseen. In the first round, player $\boldsymbol{x}$ can bet or check. Then, if player $\boldsymbol{x}$ bets player $\boldsymbol{y}$ can choose to call or fold. If player $\boldsymbol{x}$ checks then player $\boldsymbol{y}$ can bet or check. Finally, if player $\boldsymbol{x}$ checks and player $\boldsymbol{y}$ bets, then player $\boldsymbol{x}$ has a final round where they can call or fold. If neither player folded, then the player with the higher card wins the pot. If a player folded then the other player wins the pot.

We show a game tree of Kuhn poker in Figure 2. Players' imperfect information is modeled by information sets. In Figure 2, nodes with the same color belong to the same information set. A player cannot distinguish among nodes in a given information set (that belongs to this player). For example, player $\boldsymbol{x}$ cannot distinguish among the blue nodes, since in both nodes, player $\boldsymbol{x}$ was dealt Queen but does not know whether player $\boldsymbol{y}$ was dealt Jack or King.

We can further separate the decision spaces and consider them individually on a per-player basis, which is where the concept of a treeplex arises. We show player $\boldsymbol{x}$'s decision space in Figure 3. Here each circular node is an information set, which is a decision node for player $\boldsymbol{x}$ where they choose an action. For example, the blue node $\boldsymbol{h}_3$ corresponds to the initial state where player $\boldsymbol{x}$ is dealt Queen, and player $\boldsymbol{x}$ can choose to bet or check at this node. Each square node is an observation node, where player $\boldsymbol{x}$ does not make a decision, but the environment or player $\boldsymbol{y}$ makes decisions which determine the next decision node for player $\boldsymbol{x}$. Each triangular node is a terminal node, where the game ends.

Each index of treeplex $\mathcal{X}$ corresponds to a solid, directed, edge in the figure. In other words, each index corresponds to an action in finite action set Ω_h for every h in $\mathcal{H}^{\mathcal{X}}$, the set of information sets that belongs to player $\boldsymbol{x}$. Indices (solid edges) are labeled from x_1 to x_{12}. More specifically, $\boldsymbol{x} \in \mathcal{X}$ if $\boldsymbol{x}$ satisfies $x_i \geq 0$ for every index i and for every $h \in \mathcal{H}^{\mathcal{X}}$,

$$\sum_{i \in \Omega_h} x_i = x_{\sigma(h)},$$

where index $\sigma(h) \in \Omega_{h'}$ is the unique action such that h can be reached immediately by taking $\sigma(h)$ when player $\boldsymbol{x}$ is in information set h'. When no such action exists, that is, h can be reached immediately in the beginning, we set $\sigma(h) = 0$ and $x_0 = 1$. For example, we must have $x_6 = x_7 + x_8$ and $x_5 + x_6 = x_0 = 1$ if $\boldsymbol{x} \in \mathcal{X}$. Intuitively, x_i is the probability taking action i, given that the sequential decisions from the environment and player $\boldsymbol{y}$ can lead to the information set where action i is. Similarly, we show player $\boldsymbol{y}$'s decision space in Figure 4, which illustrates treeplex $\mathcal{Y}$.

B Omitted Details of Section 7

In this section, we provide more details about the experiments.

B.1 Additional Experiments

As we mentioned in Section 7, although VOGDA and DOMWU have linear convergence rate in theory, we use much larger step sizes η in the Leduc poker experiment than what Corollary 5 and Theorem 7 suggest, which explains why we were not able to observe the linear convergence. Here, we rerun this experiment with a smaller step size for VOGDA and DOMWU. With more iterations, on the order of 10^5, we observe again that they exhibit fast convergence, as shown in Figure 5.

B.2 Description of the Games

We briefly introduce the games in the experiments. Beside the rules of the games, we show the game size by providing $M, N, |\mathcal{H}^{\mathcal{Y}}|$, and $|\mathcal{H}^{\mathcal{Y}}|$ (recall $\boldsymbol{G} = \mathbb{R}^{M \times N}$).

Kuhn poker Introduced in [Kuhn, 1950], the deck for Kuhn poker contains three playing cards: King, Queen, and Jack. Each player is dealt one card, while the third card is unseen. Then a betting process proceeds. Player $\boldsymbol{x}$ can check or raise, and then player $\boldsymbol{y}$ can also check or raise. Player $\boldsymbol{x}$ has a final round to call or fold if player $\boldsymbol{x}$ checks but player $\boldsymbol{y}$ raises in the previous round. The player with the higher card wins the pot. In this game, $M = N = 13$, $|\mathcal{H}^{\mathcal{X}}| = |\mathcal{H}^{\mathcal{Y}}| = 6$.

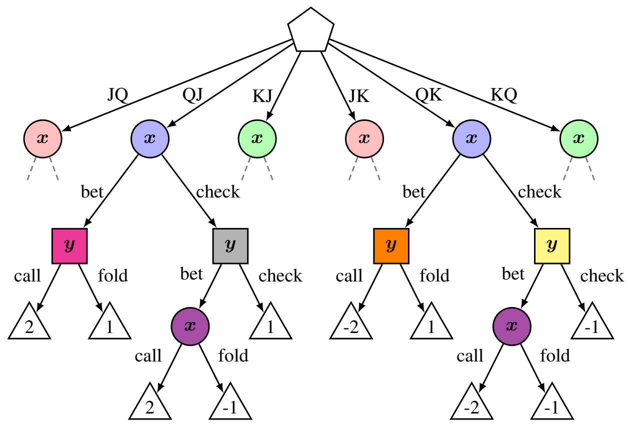

Figure 2: A game tree for Kuhn poker. The edge labeled with JQ means that player $\boldsymbol{x}$ is dealt Jack and player $\boldsymbol{y}$ is dealt Queen. The cases are similar at other edges. We omit branches stemming from the green and red nodes, which are similar to what we present for the blue nodes. Circular nodes are player $\boldsymbol{x}$'s decision nodes, while square nodes are player $\boldsymbol{y}$'s decision nodes. Triangular nodes are terminal nodes, where the values denote the utility for player $\boldsymbol{x}$ (and thus the loss for player $\boldsymbol{y}$). Nodes with the same color belong to the same information set, and a player cannot distinguish among nodes within the same information set, that is, they only know they are at one of these nodes but do not know which node they are at exactly.

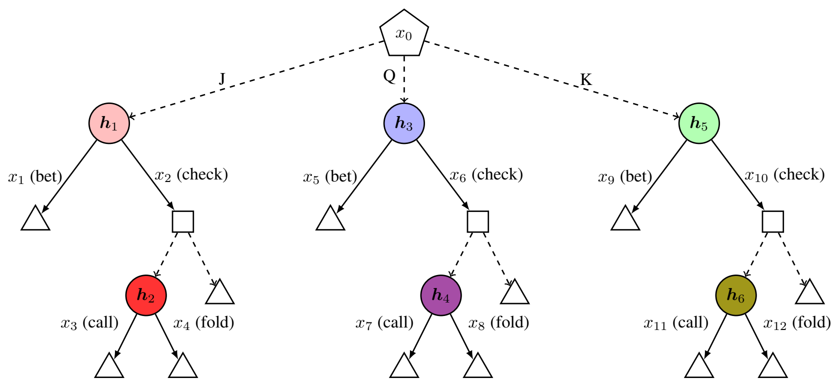

Figure 3: The decision space for player $\boldsymbol{x}$ and treeplex $\mathcal{X}$. Each circular node is an information set, each square node is an observation node, and each triangular node is a termination node, where the decision process ends. Each solid edge corresponds an action and one of the indexes in treeplex $\mathcal{X}$. More specifically, we have $M = 13$, $\mathcal{H}^{\mathcal{X}} = \{\boldsymbol{h}_1, \dots, \boldsymbol{h}_6\}$, $\Omega_{\boldsymbol{h}_i} = \{2i-1, 2i\}$ for $i = 1, \dots, 6$, $\sigma(\boldsymbol{h}_1) = \sigma(\boldsymbol{h}_3) = \sigma(\boldsymbol{h}_5) = 0$, $\sigma(\boldsymbol{h}_2) = 2$, $\sigma(\boldsymbol{h}_4) = 6$, $\sigma(\boldsymbol{h}_6) = 10$.

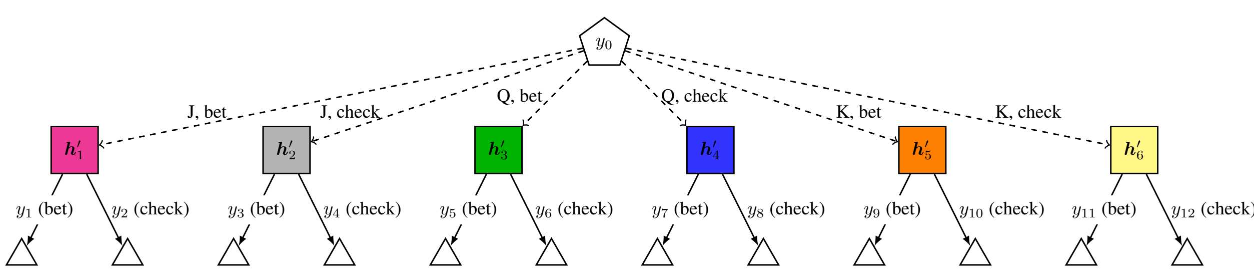

Figure 4: The decision space for player $\boldsymbol{y}$ and treeplex $\mathcal{Y}$. Each square node is an information set and each triangular node is a termination node. More specifically, we have $N = 13$, $\mathcal{H}^{\mathcal{Y}} = \{\boldsymbol{h}'_1, \dots, \boldsymbol{h}'_6\}$, $\Omega_{\boldsymbol{h}'_i} = \{2i-1, 2i\}$ and $\sigma(\boldsymbol{h}'_i) = 0$ for $i = 1, \dots, 6$.

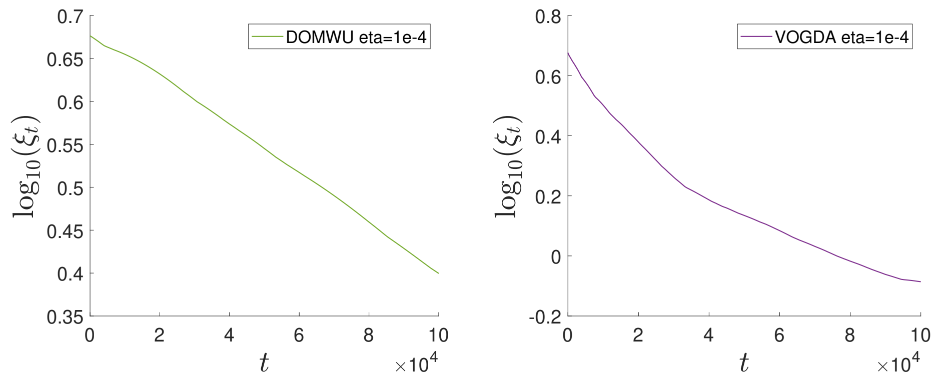

Figure 5: Experiments on Kuhn Leduc poker for DOMWU (left) and VOGDA (right) with small step sizes and more time steps.

Pursuit-evasion This is a search game considered in [Kroer et al., 2018]. Given a direct graph, player $\boldsymbol{x}$ controls an attacker to move in the graph, while player $\boldsymbol{y}$ controls two patrols, who are only allowed to move in their patrol areas. The game has 4 rounds. In each round, the attacker and the patrols act simultaneously. The attacker can move to any adjacent node or choose to wait in each round, with the final goal of going across the patrol areas and reach the goal nodes without being caught by the patrols. On the other hand, player $\boldsymbol{y}$'s goal is to let one of the patrols to reach the same node as the attacker in any round. A patrol who visits a node that was previously visited by the attacker will know that the attacker was there if the attacker did not wait at that node (in order to clean up their traces). In this game, $M = 52$, $N = 2029$, $|\mathcal{H}^{\mathcal{X}}| = 34$, $|\mathcal{H}^{\mathcal{Y}}| = 348$.

Leduc poker Introduced in [Southey et al., 2005], the game is similar to Kuhn poker. There are total 6 cards in the deck with two Kings, two Queens, and two Jacks. Each player is dealt a private card while there is another unrevealed public card. In the first round, player $\boldsymbol{y}$ bets after player $\boldsymbol{x}$ bets. After that the public card will be revealed and there is another betting stage. In a showdown stage, a player who has the same rank with the public card wins. Otherwise, the player with the higher card wins. In this game, $M = N = 337$, $|\mathcal{H}^{\mathcal{X}}| = |\mathcal{H}^{\mathcal{Y}}| = 144$.

B.3 Parameter Selection in the Dilated Regularizers

We note that for the dilated regularizers, we use the unweighted version in the experiments. This is sufficient as Lemma 9 shows there always exists an assignment $\boldsymbol{\alpha}$ such that $\psi^{\text{dil}}_{\boldsymbol{\alpha}}$ is 1-strongly convex and $\boldsymbol{\alpha} = \beta \cdot \mathbf{1}$ for some $\beta > 0$, where $\mathbf{1}$ is the all-one vector over $\mathcal{H}^{\mathcal{Z}}$. In this way, the value of η we refer to in Figure 1 is actually η/β. However, as we mentioned, the final η used in the experiments may still be larger than what the theorems suggest. Showing similar results while allowing larger η is an interesting future direction.

Lemma 9. *For an assignment $\boldsymbol{\alpha}$ such that $\psi^{dil}_{\boldsymbol{\alpha}}$ is 1-strongly convex, $\psi^{dil}_{\boldsymbol{\alpha}'}$ is also 1-strongly convex where $\boldsymbol{\alpha}' = \|\boldsymbol{\alpha}\|_\infty \mathbf{1}$.*

Proof. Recall the definition of regularizer $\psi^{\text{dil}}_{\boldsymbol{\alpha}}(\boldsymbol{z})$ in Eq. (3). Since each term $z_{\sigma(h)} \cdot \psi\left(\frac{\boldsymbol{z}_h}{z_{\sigma(h)}}\right)$ is convex in $\boldsymbol{z}$ (thus with a non-negative Bregman divergence), $D_{\psi^{\text{dil}}_{\boldsymbol{\alpha}}}$ is an increasing function in any variable α_h ($h \in \mathcal{H}^{\mathcal{Z}}$). Therefore, we have

$$\frac{\|\boldsymbol{z} - \boldsymbol{z}'\|^2}{2} \leq D_{\psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}, \boldsymbol{z}') \leq D_{\psi^{\text{dil}}_{\boldsymbol{\alpha}'}}(\boldsymbol{z}, \boldsymbol{z}'),$$

which completes the proof. □

C Omitted Details of Section 4

When introducing VOMWU in Section 4, we mention that Ψ^{van} is 1-strongly convex with respect to the 2-norm. In the following, we formally show this result. Before that, we first show a technical lemma.

Lemma 10. *For $u, v \in [0, 1]$, the following inequalities hold:*

$$\frac{(u-v)^2}{2} \leq u \ln\left(\frac{u}{v}\right) - u + v \leq \frac{(u-v)^2}{v}. \tag{14}$$

Proof. Define $f(u,v) = u\ln\left(\frac{u}{v}\right) - u + v - \frac{(u-v)^2}{2}$ and $g(u,v) = \frac{(u-v)^2}{v} - u\ln\left(\frac{u}{v}\right) + u - v$. To prove the claim, it is sufficient to show that the minimum of each of the two functions is zero. Since both functions have the only critical point $(0,0)$, there is no extreme point in the interior. Also, it is straightforward to find a (u,v) such that $f(u,v), g(u,v) > 0$ in the interior. Thus, it remains to check if the boundary of the domain satisfies $f(u,v), g(u,v) \geq 0$. For $u = 0$, we have

$$\frac{(0-v)^2}{2} = \frac{v^2}{2} \leq v = \frac{(0-v)^2}{v}.$$

The case for $v = 0$ is trivial. For $u = 1$, note that

$$\frac{(v-1)^2}{2} \leq v - 1 - \ln(v) \leq \frac{(v-1)^2}{v}$$

when $0 \leq v \leq 1$. For $v = 1$, we have

$$\frac{(u-1)^2}{2} \leq u \ln u - u + 1 \leq (u-1)^2$$

when $0 \leq u \leq 1$. Therefore, we conclude $f(u,v), g(u,v) \geq 0$ and this finishes the proof. □

Now we are ready to give the result.

Lemma 11. *Ψ^{van} is 1-strongly convex with respect to the 2-norm.*

Proof. The result follows by the first inequality of Eq. (14) and $0 \leq z_i \leq 1$ for all $\boldsymbol{z} \in \mathcal{Z}$. □

D CFR-based Algorithms and Proof of Theorem 3

In this section, we first introduce CFR, CFR+, and their optimistic versions. Then we show Theorem 3 in Appendix D.3 and their empirical last-iterate divergence in Appendix D.4.

D.1 CFR and Its Optimistic Version

Given P-dimensional treeplex $\mathcal{Z}$, loss vector $\boldsymbol{\ell} \in \mathbb{R}^P$, and $\boldsymbol{z} \in \mathcal{Z}$, we recursively define the value vector $\boldsymbol{L} \in \mathbb{R}^P$, to be

$$L_i = \ell_i + \sum_{g \in \mathcal{H}_i} \langle \boldsymbol{q}_g, \boldsymbol{L}_g \rangle$$

for every index i (recall that $q_i = z_i / z_{p_i}$). At time t, given $\boldsymbol{q}_t$, loss vector ℓ_t, and its value vector $\boldsymbol{L}_t$, denote $\text{reg}^g_{t,j} = \langle \boldsymbol{q}_{t,g}, \boldsymbol{L}_{t,g} \rangle - L_{t,j}$, $\text{Reg}^g_{0,j} = 0$, and $\text{Reg}^g_{t,j} = \text{Reg}^g_{t-1,j} + \text{reg}^g_{t,j}$ for every simplex $g \in \mathcal{H}^{\mathcal{Z}}$ and index $j \in \Omega_g$. In the literature, *Counterfactual Regret Minimization* (CFR) Zinkevich et al. [2007] refers to the algorithm running *regret matching* on every simplex in a treeplex. Specifically, on simplex g at time $t+1$, regret matching plays arbitrarily when $\sum_{j \in \Omega_g} \left[\text{Reg}^g_{t,j}\right]_+ = 0$, where $[x]_+ = \max(0, x)$; otherwise, it plays

$$q_{t+1,i} = \frac{\left[\text{Reg}^g_{t,i}\right]_+}{\sum_{j \in \Omega_g} \left[\text{Reg}^g_{t,j}\right]_+},$$

for all i in Ω_g. In the two-player zero-sum setting, player $\boldsymbol{x}$ runs CFR on every simplex in treeplex $\mathcal{X}$ along with point $\boldsymbol{x}_t \in \mathcal{X}$ and loss vector $\boldsymbol{L}_t^{\boldsymbol{x}} = \boldsymbol{G}\boldsymbol{y}_t$, and player $\boldsymbol{y}$ runs CFR on every simplex in $\mathcal{Y}$ along with point $\boldsymbol{y}_t \in \mathcal{Y}$ and loss vector $\boldsymbol{L}_t^{\boldsymbol{y}} = -\boldsymbol{G}^\top \boldsymbol{x}_t$ for every time t.

The optimistic version of CFR Farina et al. [2021a] is running the optimistic version of regret matching on every simplex in a treeplex. Specifically, the algorithm plays arbitrarily when $\sum_{j\in\Omega_g}\left[\mathrm{Reg}_{t,j}^g + \mathrm{reg}_{t,j}^g\right]_+ = 0$; otherwise,

$$q_{t+1,i} = \frac{\left[\mathrm{Reg}_{t,i}^g + \mathrm{reg}_{t,i}^g\right]_+}{\sum_{j\in\Omega_g}\left[\mathrm{Reg}_{t,j}^g + \mathrm{reg}_{t,j}^g\right]_+}.$$

To get an approximate Nash equilibrium at time t, CFR and its optimistic version consider the average iterate, that is, they return

$$(\overline{\boldsymbol{x}}_t, \overline{\boldsymbol{y}}_t) = \left(\frac{1}{t}\sum_{\tau=1}^{t}\boldsymbol{x}_\tau, \frac{1}{t}\sum_{\tau=1}^{t}\boldsymbol{y}_\tau\right).$$

D.2 CFR+ and Its Optimistic Version

To introduce CFR+, we first introduce another regret-minimization algorithm on simplex, *regret matching+*. Similar to $\mathrm{Reg}_{t,j}^g$, we define $\widehat{\mathrm{Reg}}_{0,j}^g = 0$ and

$$\widehat{\mathrm{Reg}}_{t,j}^g = \left[\widehat{\mathrm{Reg}}_{t-1,j}^g + \mathrm{reg}_{t,j}^g\right]_+,$$

for every simplex $g \in \mathcal{H}^{\mathcal{Z}}$ and index $j \in \Omega_g$. On simplex g at time $t+1$, regret matching+ plays arbitrarily when $\sum_{j\in\Omega_g}\left[\widehat{\mathrm{Reg}}_{t,j}^g\right]_+ = 0$; otherwise, it plays

$$q_{t+1,i} = \frac{\left[\widehat{\mathrm{Reg}}_{t,i}^g\right]_+}{\sum_{j\in\Omega_g}\left[\widehat{\mathrm{Reg}}_{t,j}^g\right]_+}.$$

CFR+ Tammelin [2014] refers to running *regret matching+* on every simplex in a treeplex. In the two-player zero-sum setting, CFR+ usually refers to the one with *alternating updates*. Specifically, player $\boldsymbol{x}$ runs CFR+ on every simplex in $\mathcal{H}^{\mathcal{X}}$ with loss vector $\boldsymbol{L}_t^{\boldsymbol{x}} = \boldsymbol{G}\boldsymbol{y}_t$ and player $\boldsymbol{y}$ runs CFR on every simplex in $\mathcal{H}^{\mathcal{Y}}$ with loss vector $\boldsymbol{L}_t^{\boldsymbol{y}} = -\boldsymbol{G}^\top\boldsymbol{x}_{t+1}$ (note that in the case with simultaneous updates, $\boldsymbol{L}_t^{\boldsymbol{y}} = -\boldsymbol{G}^\top\boldsymbol{x}_t$).

The optimistic version of CFR+ Farina et al. [2021a] is running the optimistic version of regret matching+ on every simplex in a treeplex. Specifically, the algorithm plays arbitrarily when $\sum_{j\in\Omega_g}\left[\widehat{\mathrm{Reg}}_{t,j}^g + \mathrm{reg}_{t,j}^g\right]_+ = 0$; otherwise, it plays

$$q_{t+1,i} = \frac{\left[\widehat{\mathrm{Reg}}_{t,i}^g + \mathrm{reg}_{t,i}^g\right]_+}{\sum_{j\in\Omega_g}\left[\widehat{\mathrm{Reg}}_{t,j}^g + \mathrm{reg}_{t,j}^g\right]_+}.$$

Regarding the averaging scheme, CFR+ usually refers to the version with *linear averaging* to get an approximate Nash equilibrium at time t. Specifically, it returns

$$(\overline{\boldsymbol{x}}_t, \overline{\boldsymbol{y}}_t) = \left(\frac{2}{t(t+1)}\sum_{\tau=1}^{t}\tau\cdot\boldsymbol{x}_\tau, \frac{2}{t(t+1)}\sum_{\tau=1}^{t}\tau\cdot\boldsymbol{y}_\tau\right).$$

In summary, CFR+ and its optimistic version refer to running regret matching+ and optimistic regret matching+ on every simplex with alternating updates and linear averaging.

D.3 Proof of Theorem 3

Proof of Theorem 3. Note that in this instance, there is only one simplex $g^{\boldsymbol{x}}$ for player $\boldsymbol{x}$ and one simplex $g^{\boldsymbol{y}}$ for player $\boldsymbol{y}$. The game matrix $\boldsymbol{G}$ of the rock-paper-scissors is

$$\boldsymbol{G} = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix} = -\boldsymbol{G}^\top.$$

We consider the case when $\boldsymbol{x}_1 = \boldsymbol{y}_1$. Recall that we have $\boldsymbol{L}_t^{\boldsymbol{x}} = \boldsymbol{G}\boldsymbol{y}_t$ and $\boldsymbol{L}_t^{\boldsymbol{y}} = -\boldsymbol{G}^\top \boldsymbol{x}_t$. Therefore, we know that

$$\boldsymbol{L}_1^{\boldsymbol{x}} = \boldsymbol{G}^\top \boldsymbol{y}_1 = \boldsymbol{G}^\top \boldsymbol{x}_1 = -\boldsymbol{G}^\top \boldsymbol{x}_1 = \boldsymbol{L}_1^{\boldsymbol{y}},$$

and

$$\text{reg}_{1,j}^{g^{\boldsymbol{x}}} = \boldsymbol{x}_1^\top \boldsymbol{L}_1^{\boldsymbol{x}} - L_{1,j}^{\boldsymbol{x}} = \boldsymbol{x}_1^\top \boldsymbol{G}\boldsymbol{x}_1 - L_{1,j}^{\boldsymbol{x}} = -L_{1,j}^{\boldsymbol{x}} = -L_{1,j}^{\boldsymbol{y}} = \text{reg}_{1,j}^{g^{\boldsymbol{y}}}$$

for $j = 1, 2, 3$. It is not hard to see that in this case, we have $\boldsymbol{x}_2 = \boldsymbol{y}_2$, and thus $\boldsymbol{x}_t = \boldsymbol{y}_t$ for every t. Consequently, it is sufficient to focus on the updates of $\boldsymbol{x}_t$. For notional convenience, we write

$$\text{reg}_t = \text{reg}_t^{g^{\boldsymbol{x}}} = \boldsymbol{G}^\top \boldsymbol{x}_t = (x_{t,2} - x_{t,3}, x_{t,3} - x_{t,1}, x_{t,1} - x_{t,2})^\top, \tag{15}$$

$$\text{Reg}_t = \text{Reg}_t^{g^{\boldsymbol{x}}} = \text{Reg}_{t-1}^{g^{\boldsymbol{x}}} + \text{reg}_t = \sum_{\tau=1}^{t} \text{reg}_\tau, \tag{16}$$

and thus

$$x_{t+1,j} = \frac{\left[\text{Reg}_{t,j}\right]_+}{\left[\text{Reg}_{t,1}\right]_+ + \left[\text{reg}_{t,2}\right]_+ + \left[\text{Reg}_{t,3}\right]_+} \tag{17}$$

for $j = 1, 2, 3$. We call distribution $\boldsymbol{x}_t$ *imbalanced* if there exists a permutation λ of $\{1, 2, 3\}$ such that

$$x_{t,\lambda(1)} \geq x_{t,\lambda(2)} \geq 0 = x_{t,\lambda(3)}.$$

We prove that if $\boldsymbol{x}_1$ is imbalanced, then every $\boldsymbol{x}_t$ is imbalanced. Suppose at some time t, $\boldsymbol{x}_t$ is imbalanced and the corresponding λ is the identity without loss of generality, that is,

$$x_{t,1} \geq x_{t,2} \geq 0 = x_{t,3}. \tag{18}$$

In this case, we know that $\text{Reg}_{t-1,1} > 0$. By Eq. (18) and Eq. (15), we also know that $\text{reg}_{t,1} \geq 0$, $\text{reg}_{t,3} \geq 0$, $\text{reg}_{t,2} < 0$, and $\text{Reg}_{t,1} > 0$. Moreover, by Eq. (15) and Eq. (16), we can get

$$\text{reg}_{t,1} + \text{reg}_{t,2} + \text{reg}_{t,3} = 0, \ \text{Reg}_{t,1} + \text{Reg}_{t,2} + \text{Reg}_{t,3} = 0.$$

Therefore, we have $\text{Reg}_{t,2} + \text{Reg}_{t,3} < 0$, which means that at least one of $x_{t+1,2}$ and $x_{t+1,3}$ is zero. Moreover, we have

$$\text{Reg}_{t,1} = \text{Reg}_{t-1,1} + \text{reg}_{t,1} \geq \text{Reg}_{t-1,1} \geq \text{Reg}_{t-1,2} > \text{Reg}_{t-1,2} + \text{reg}_{t,2} = \text{Reg}_{t,2},$$

where the second inequality follows from $x_{t,1} \geq x_{t,2}$ and Eq. (17). The inequalities above imply that $x_{t+1,1} > x_{t+1,2}$. Thus, we get one of the following three situations continues to hold at time $t+1$:

$$x_{t+1,1} \geq x_{t+1,2} \geq 0 = x_{t+1,3}, \tag{19}$$

$$x_{t+1,1} \geq x_{t+1,3} \geq 0 = x_{t+1,2}, \tag{20}$$

$$x_{t+1,3} \geq x_{t+1,1} \geq 0 = x_{t+1,2}. \tag{21}$$

If Eq. (19) holds at time $t+1$, the same argument implies one of the three arguments above continues to hold; otherwise, if at some time step $\tau > t$, Eq. (20) holds, that is,

$$x_{\tau,1} \geq x_{\tau,3} \geq 0 = x_{\tau,2}. \tag{22}$$

Similarly, we know that $\text{reg}_{\tau,1} \leq 0, \text{reg}_{\tau,2} \leq 0$ and $\text{reg}_{\tau,3} \geq 0$, and $x_{\tau+1,2} = 0$. Thus, we get either Eq. (22) continues to hold at time $\tau + 1$ or

$$x_{\tau+1,3} \geq x_{\tau+1,1} \geq 0 = x_{\tau+1,2}$$

holds, which is exactly the same permutation in Eq. (21). With similar arguments, we know that for every imbalanced distribution, either the same permutation holds in the next round, or it transits to another imbalanced distribution with another permutation. Note that the average iterate of the sequence $\{\boldsymbol{x}_t\}_t$ converges to the uniform distribution as CFR is a no-regret algorithm, so $\boldsymbol{x}_t$ never converges to any imbalanced distribution. Therefore, we conclude that $\boldsymbol{x}_t$ diverges if the algorithm starts from $x_1 = y_1$ being an arbitrary imbalanced distribution. □

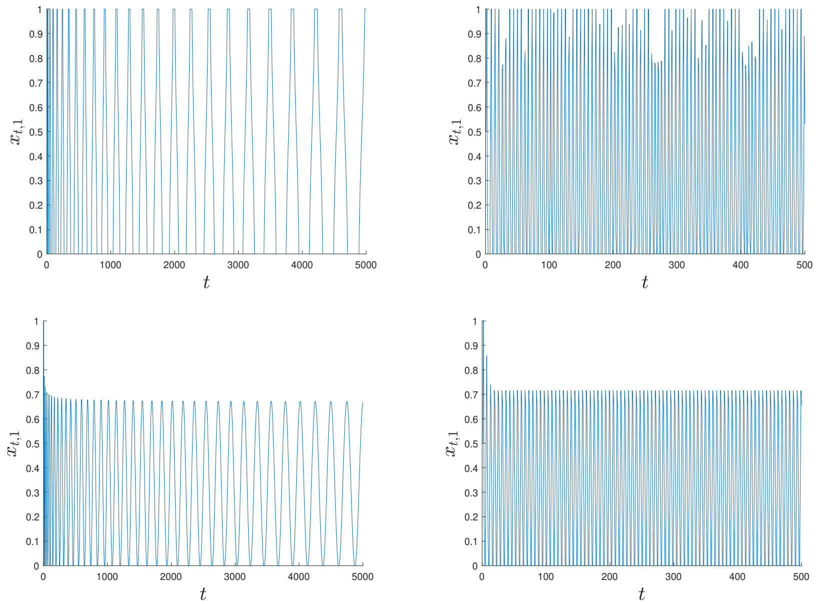

Figure 6: Last-iterate divergence of CFR, CFR+ with simultaneous updates, and their optimistic versions in the rock-paper-scissors game. The first row is the CFR algorithms while the second row is the CFR+ algorithms. For both rows, we put the vanilla version on the left and the optimistic version on the right. All algorithms start with $\boldsymbol{x}_1 = \boldsymbol{y}_1 = (1, 0, 0)^\top$.

D.4 Experiments

Besides Theorem 3, we empirically observe divergence of CFR, CFR+ with simultaneous updates, and their optimistic versions in the rock-paper-scissors game. The results are shown in Figure 6. We remark that in these experiments, we consider CFR+ with simultaneous updates instead of the more commonly used ones (alternating updates). In fact, we observe that with alternating updates, the optimistic CFR+ empirically has last-iterate convergence in some matrix games. As all of our theoretical results are for simultaneous updates, the theoretical justification of this observation is beyond the scope of this paper, but it is an interesting direction for future works.

E Proofs of Theorem 6 and Theorem 7

In this section, we show the proof of Theorem 6 in Appendix E.4 and the proof of Theorem 7 in Appendix E.5. We generally follow the outline in Section 6.3. We discuss the technical difficulty to get a convergence rate for DOGDA in Appendix E.6. Throughout the section, we assume that $\boldsymbol{G}$ has a unique Nash equilibrium $\boldsymbol{z}^* = (\boldsymbol{x}^*, \boldsymbol{y}^*)$ (call it the uniqueness assumption). For the sake of analysis, we equivalently define $\Psi^{\text{dil}}_{\boldsymbol{\alpha}}$ as

$$\Psi^{\text{dil}}_{\boldsymbol{\alpha}}(\boldsymbol{z}) = \sum_i \alpha_{h(i)} z_i \ln \frac{z_i}{\sum_{j \in \Omega_{h(i)}} z_j}.$$

Note that under this definition, we have

$$\frac{\partial \Psi^{\text{dil}}_{\boldsymbol{\alpha}}(\boldsymbol{z})}{\partial z_i} = \alpha_{h(i)} \left[\ln \left(\frac{z_i}{\sum_{j \in \Omega_{h(i)}} z_j} \right) + 1 - \sum_{k \in \Omega_{h(i)}} \frac{z_k}{\sum_{j \in \Omega_{h(i)}} z_j} \right] = \alpha_{h(i)} \ln q_i, \tag{23}$$

and

$$D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}, \boldsymbol{z}') = \sum_i \alpha_{h(i)} \left[z_i \ln(q_i) - z'_i \ln(q'_i) - (z_i - z'_i) \ln(q'_i) \right] = \sum_i \alpha_{h(i)} z_i \ln \frac{q_i}{q'_i}. \tag{24}$$

E.1 Strict Complementary Slackness and Proof of Eq. (6) in EFGs

In this subsection, we prove Eq. (6), which is restated in Lemma 12.

Lemma 12. *Under the uniqueness assumption, for both* VOMWU *and* DOMWU*, there exists some constant* $C_{12} > 0$ *that depends on the game,* η*, and* $\widehat{\boldsymbol{z}}_1$ *such that for any* t*,*

$$\zeta_t = D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) + D_\psi(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t) \geq C_{12}\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_{t+1}\|^2. \tag{25}$$

Before proving this lemma, we show some useful properties of EFGs. The first one is the strict complementary slackness.

E.1.1 Strict Complementary Slackness

We formulate the minimax problem in Eq. (1) as a linear program. This is a standard procedure in the literature (see, for example, [Nisan et al., 2007, Section 3.11]), but we show its derivation here for completeness. Based on Definition 1, we have $\boldsymbol{x} \in \mathcal{X}$ if $\boldsymbol{x}$ satisfies $x_i \geq 0$ for every i and

$$x_0 = 1, \qquad \sum_{i \in \Omega_h} x_i = x_{\sigma(h)}, \; \forall h \in \mathcal{H}^{\mathcal{X}}. \tag{26}$$

We can write the constraints in Eq. (26) using a matrix $\boldsymbol{A} \in \mathbb{R}^{(|\mathcal{H}^{\mathcal{X}}|+1)\times M}$ such that $\boldsymbol{x} \in \mathcal{X}$ if and only if $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{e}, \; \boldsymbol{x} \geq \boldsymbol{0}$, where all entries in $\boldsymbol{e}$ are zero except for $e_0 = 1$, which corresponds to the constraint $x_0 = e_0 = 1$. Similarly, for player $\boldsymbol{y}$, we have the constraint matrix $\boldsymbol{B}$ and vector $\boldsymbol{f}$. Consequently, we write the best response of $\boldsymbol{x}$ to a fixed $\boldsymbol{y}$ to as the following linear program:

$$\min_{\boldsymbol{x}} \boldsymbol{x}^\top(\boldsymbol{G}\boldsymbol{y}), \text{ subject to } \boldsymbol{A}\boldsymbol{x} = \boldsymbol{e}, \; \boldsymbol{x} \geq \boldsymbol{0}.$$

The dual of this linear program is

$$\max_{\boldsymbol{V}} \boldsymbol{e}^\top \boldsymbol{V}, \text{ subject to } \boldsymbol{A}^\top \boldsymbol{V} \leq \boldsymbol{G}\boldsymbol{y},$$

where $\boldsymbol{V} \in \mathbb{R}^{|\mathcal{H}^{\mathcal{X}}|+1}$. Recall that all entries in $\boldsymbol{e}$ are zero except for $e_0 = 1$ so the objective $\boldsymbol{e}^\top \boldsymbol{V} = V_0$. On the other hand, player $\boldsymbol{y}$ tries to maximize $\boldsymbol{x}^\top \boldsymbol{G}\boldsymbol{y}$, that is, maximize V_0 by the strong duality. Therefore, every $\boldsymbol{y}^* \in \mathcal{Y}^*$ is a solution to the following maximin problem

$$\max_{\boldsymbol{V},\boldsymbol{y}} V_0, \text{ subject to } \boldsymbol{A}^\top \boldsymbol{V} \leq \boldsymbol{G}\boldsymbol{y}, \; \boldsymbol{B}\boldsymbol{y} = \boldsymbol{f}, \; \boldsymbol{y} \geq \boldsymbol{0}, \tag{27}$$

which is also a linear program. The dual of this linear problem is

$$\min_{\boldsymbol{U},\boldsymbol{x}} U_0, \text{ subject to } \boldsymbol{B}^\top \boldsymbol{U} \geq \boldsymbol{G}^\top \boldsymbol{x}, \; \boldsymbol{A}\boldsymbol{x} = \boldsymbol{e}, \; \boldsymbol{x} \geq \boldsymbol{0}. \tag{28}$$

It is also not hard to see that every $\boldsymbol{x}^* \in \mathcal{X}$ is a solution to this dual. We conclude that Eq. (27) and Eq. (28) are primal-dual linear programs of the minmax problem in Eq. (1). Given an optimal solution pair $\boldsymbol{x}^*, \boldsymbol{y}^*$ along with $\boldsymbol{V}^*, \boldsymbol{U}^*$, by the complementary slackness, we have some slackness variables $\boldsymbol{w}^* \in \mathbb{R}^M, \boldsymbol{s}^* \in \mathbb{R}^N$ such that $\boldsymbol{x}^* \odot \boldsymbol{w}^* = \boldsymbol{0}, \boldsymbol{y}^* \odot \boldsymbol{s}^* = \boldsymbol{0}$ ($\odot$ denotes the element-wise product), and

$$\boldsymbol{A}^\top \boldsymbol{V}^* - \boldsymbol{G}\boldsymbol{y}^* + \boldsymbol{w}^* = \boldsymbol{0}, \quad \boldsymbol{B}^\top \boldsymbol{U}^* - \boldsymbol{G}^\top \boldsymbol{x}^* - \boldsymbol{s}^* = \boldsymbol{0}, \quad \boldsymbol{w}^*, \boldsymbol{s}^* \geq \boldsymbol{0}. \tag{29}$$

Thus, Eq. (29) implies that for every index i of player $\boldsymbol{x}$,

$$V^*_{h(i)} - \sum_{g \in \mathcal{H}_i} V^*_g - (\boldsymbol{G}\boldsymbol{y}^*)_i = (\boldsymbol{A}^\top \boldsymbol{V}^*)_i - (\boldsymbol{G}\boldsymbol{y}^*)_i = (\boldsymbol{A}^\top \boldsymbol{V}^* - \boldsymbol{G}\boldsymbol{y}^*)_i = -w^*_i. \tag{30}$$

Additionally, the *strict complementary slackness* (see, for example, [Vanderbei et al., 2015, Theorem 10.7]) ensures that there exists an optimal solution $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ such that

$$\boldsymbol{x}^* + \boldsymbol{w}^* > \boldsymbol{0}, \quad \boldsymbol{y}^* + \boldsymbol{s}^* > \boldsymbol{0}. \tag{31}$$

Under the uniqueness assumption, the strict complementary slackness must hold for the unique optimal solution $(\boldsymbol{x}^*, \boldsymbol{y}^*)$. Therefore, when $x^*_i > 0$, we have $w^*_i = 0$, which means that Eq. (30) implies

$$V^*_h = (\boldsymbol{G}\boldsymbol{y}^*)_i + \sum_{g \in \mathcal{H}_i} V^*_g, \quad \forall h \in \mathcal{H}^{\mathcal{X}}, \; i \in \Omega_h, \; x^*_i > 0; \tag{32}$$

otherwise, if $x_i^*=0$ and $w_i^*>0$, by Eq. (30), we have

$$V_h^* < (\boldsymbol{G}\boldsymbol{y}^*)_i + \sum_{g\in\mathcal{H}_i} V_g^*, \quad \forall h\in\mathcal{H}^{\mathcal{X}},\ i\in\Omega_h,\ x_i^*=0. \tag{33}$$

Note that for every terminal index i, $\mathcal{H}_i$ is empty and we set the term $\sum_{g\in\mathcal{H}_i} V_g^*$ zero correspondingly. The case is similar for $\boldsymbol{U}^*$ and $\boldsymbol{G}^\top\boldsymbol{x}^*$. We summarize this result as the following lemma.

Lemma 13. *Under the uniqueness assumption, we have*

$$\begin{aligned}
&\sum_{g\in\mathcal{H}_i} V_g^* + (\boldsymbol{G}\boldsymbol{y}^*)_i = V_{h(i)}^* && \forall i\in \mathit{supp}(\boldsymbol{x}^*),\\
&\sum_{g\in\mathcal{H}_i} V_g^* + (\boldsymbol{G}\boldsymbol{y}^*)_i > V_{h(i)}^* && \forall i\notin \mathit{supp}(\boldsymbol{x}^*),\\
&\sum_{g\in\mathcal{H}_j} U_g^* + (\boldsymbol{G}^\top\boldsymbol{x}^*)_j = U_{h(j)}^* && \forall j\in \mathit{supp}(\boldsymbol{y}^*),\\
&\sum_{g\in\mathcal{H}_j} U_g^* + (\boldsymbol{G}^\top\boldsymbol{x}^*)_j < U_{h(j)}^* && \forall j\notin \mathit{supp}(\boldsymbol{y}^*).
\end{aligned}$$

E.1.2 Some Problem-dependent Constants

After introducing the strict complementary slackness, we are ready to introduce some problem-dependent constants. Note that by Lemma 13, we have the following constant $\xi>0$.

Definition 2. *Under the uniqueness assumption, we define*

$$\xi \triangleq \min\left\{ \min_{i\notin \mathit{supp}(\boldsymbol{x}^*)} \sum_{g\in\mathcal{H}_i} V_g^* + (\boldsymbol{G}\boldsymbol{y}^*)_i - V_{h(i)}^*,\ \min_{j\notin \mathit{supp}(\boldsymbol{y}^*)} U_{h(j)}^* - (\boldsymbol{G}^\top\boldsymbol{x}^*)_j - \sum_{g\in\mathcal{H}_j} U_g^* \right\} \in (0,2].$$

Note that $\xi\le 2$ follows from the fact that for any information set $h\in\mathcal{H}^{\mathcal{X}}$, indices $i,j\in\Omega_h$ such that $i\notin \operatorname{supp}(\boldsymbol{x}^*)$ and $j\in\operatorname{supp}(\boldsymbol{x}^*)$, by Lemma 13, we have

$$\xi \le \sum_{g\in\mathcal{H}_i} V_g^* + (\boldsymbol{G}\boldsymbol{y}^*)_i - V_h^* = (\boldsymbol{G}\boldsymbol{y}^*)_i - (\boldsymbol{G}\boldsymbol{y}^*)_j \le 2.$$

Below, we define $\mathcal{V}^*(\mathcal{Z}) = \mathcal{V}^*(\mathcal{X})\times\mathcal{V}^*(\mathcal{Y})$, where

$$\mathcal{V}^*(\mathcal{X}) \triangleq \{\boldsymbol{x} : \boldsymbol{x}\in\mathcal{X},\ \operatorname{supp}(\boldsymbol{x})\subseteq\operatorname{supp}(\boldsymbol{x}^*)\}$$

and

$$\mathcal{V}^*(\mathcal{Y}) \triangleq \{\boldsymbol{y} : \boldsymbol{y}\in\mathcal{Y},\ \operatorname{supp}(\boldsymbol{y})\subseteq\operatorname{supp}(\boldsymbol{y}^*)\}.$$

Definition 3.

$$c_x \triangleq \min_{\boldsymbol{x}\in\mathcal{X}\backslash\{\boldsymbol{x}^*\}} \max_{\boldsymbol{y}\in\mathcal{V}^*(\mathcal{Y})} \frac{(\boldsymbol{x}-\boldsymbol{x}^*)^\top\boldsymbol{G}\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{x}^*\|_1}, \qquad c_y \triangleq \min_{\boldsymbol{y}\in\mathcal{Y}\backslash\{\boldsymbol{y}^*\}} \max_{\boldsymbol{x}\in\mathcal{V}^*(\mathcal{X})} \frac{\boldsymbol{x}^\top\boldsymbol{G}(\boldsymbol{y}^*-\boldsymbol{y})}{\|\boldsymbol{y}^*-\boldsymbol{y}\|_1}.$$

The following lemma shows that c_x and c_y are well-defined even though the outer minimization is over an open set. The proof generally follows [Wei et al., 2021, Lemma 14] but requires the results derived in Appendix E.1.1.

Lemma 14. *c_x and c_y are well-defined, and* $0<c_x,c_y\le 1$.

Proof. We first show c_x and c_y are well-defined. To simplify the notations, we define $x^*_{\min} \triangleq \min_{i\in\operatorname{supp}(\boldsymbol{x}^*)} x_i^*$ and $\mathcal{X}' \triangleq \{\boldsymbol{x} : \boldsymbol{x}\in\mathcal{X},\ \|\boldsymbol{x}-\boldsymbol{x}^*\|_1 \ge x^*_{\min}\}$, and define $y^*_{\min}$ and $\mathcal{Y}'$ similarly. We will show that

$$c_x = \min_{\boldsymbol{x}\in\mathcal{X}'} \max_{\boldsymbol{y}\in\mathcal{V}^*(\mathcal{Y})} \frac{(\boldsymbol{x}-\boldsymbol{x}^*)^\top\boldsymbol{G}\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{x}^*\|_1}, \quad c_y = \min_{\boldsymbol{y}\in\mathcal{Y}'} \max_{\boldsymbol{x}\in\mathcal{V}^*(\mathcal{X})} \frac{\boldsymbol{x}^\top\boldsymbol{G}(\boldsymbol{y}^*-\boldsymbol{y})}{\|\boldsymbol{y}^*-\boldsymbol{y}\|_1},$$

which are well-defined as the outer minimization is now over a closed set. To prove the equality for c_x, it suffices to show that for any $\boldsymbol{x} \in \mathcal{X}$ such that $\boldsymbol{x} \neq \boldsymbol{x}^*$ and $\|\boldsymbol{x} - \boldsymbol{x}^*\|_1 < x^*_{\min}$, there exists $\boldsymbol{x}' \in \mathcal{X}$ such that $\|\boldsymbol{x}' - \boldsymbol{x}^*\|_1 = x^*_{\min}$ and

$$\frac{(\boldsymbol{x} - \boldsymbol{x}^*)^\top \boldsymbol{G}\boldsymbol{y}}{\|\boldsymbol{x} - \boldsymbol{x}^*\|_1} = \frac{(\boldsymbol{x}' - \boldsymbol{x}^*)^\top \boldsymbol{G}\boldsymbol{y}}{\|\boldsymbol{x}' - \boldsymbol{x}^*\|_1}, \ \forall \boldsymbol{y}. \tag{34}$$

In fact, we can simply choose $\boldsymbol{x}' = \boldsymbol{x}^* + (\boldsymbol{x} - \boldsymbol{x}^*) \cdot \frac{x^*_{\min}}{\|\boldsymbol{x} - \boldsymbol{x}^*\|_1}$. We first argue that $\boldsymbol{x}'$ is still in $\mathcal{X}$. For each index j, if $x_j - x^*_j \geq 0$, we surely have $x'_j \geq x^*_j + 0 \geq 0$; otherwise, $x^*_j > x_j \geq 0$ and thus $j \in \text{supp}(\boldsymbol{x}^*)$ and $x^*_j \geq x^*_{\min}$, which implies $x'_j \geq x^*_j - |x_j - x^*_j| \cdot \frac{x^*_{\min}}{\|\boldsymbol{x} - \boldsymbol{x}^*\|_1} \geq x^*_j - x^*_{\min} \geq 0$. In addition, for any $h \in \mathcal{H}^{\mathcal{X}}$,

$$\begin{aligned}
\sum_{j \in \Omega_h} x'_j &= \frac{x^*_{\min}}{\|\boldsymbol{x} - \boldsymbol{x}^*\|_1} \cdot \sum_{j \in \Omega_h} x_j + \left(1 - \frac{x^*_{\min}}{\|\boldsymbol{x} - \boldsymbol{x}^*\|_1}\right) \sum_{j \in \Omega_h} x^*_j \\
&= \frac{x^*_{\min}}{\|\boldsymbol{x} - \boldsymbol{x}^*\|_1} \cdot x_{\sigma(h)} + \left(1 - \frac{x^*_{\min}}{\|\boldsymbol{x} - \boldsymbol{x}^*\|_1}\right) x^*_{\sigma(h)} \\
&= x'_{\sigma(h)}.
\end{aligned}$$

Therefore, we conclude $\boldsymbol{x}' \in \mathcal{X}$. Moreover, according to the definition of $\boldsymbol{x}'$, $\|\boldsymbol{x}' - \boldsymbol{x}^*\|_1 = x^*_{\min}$ holds. Also, since $\boldsymbol{x}^* - \boldsymbol{x}$ and $\boldsymbol{x}^* - \boldsymbol{x}'$ are parallel vectors, Eq. (34) is satisfied. The arguments above show that the c_x in Definition 3 is a well-defined real number. The case of c_y is similar.

Now we show $0 < c_x, c_y \leq 1$. The fact that $c_x, c_y \leq 1$ is a direct consequence of the definitions. Below, we use contradiction to prove that $c_y > 0$. First, if $c_y < 0$, then there exists $\boldsymbol{y} \neq \boldsymbol{y}^*$ such that ${\boldsymbol{x}^*}^\top \boldsymbol{G}\boldsymbol{y}^* < {\boldsymbol{x}^*}^\top \boldsymbol{G}\boldsymbol{y}$. This contradicts with the fact that $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is the equilibrium.

On the other hand, if $c_y = 0$, then there is some $\boldsymbol{y} \neq \boldsymbol{y}^*$ such that

$$\max_{\boldsymbol{x} \in \mathcal{V}^*(\mathcal{X})} \boldsymbol{x}^\top \boldsymbol{G}(\boldsymbol{y}^* - \boldsymbol{y}) = 0. \tag{35}$$

Consider the point $\boldsymbol{y}' = \boldsymbol{y}^* + \frac{\xi}{2N}(\boldsymbol{y} - \boldsymbol{y}^*)$ (recall the definition of ξ in Definition 2 and that $0 < \xi \leq 2$), which is a convex combination of $\boldsymbol{y}^*$ and $\boldsymbol{y}$, and hence $\boldsymbol{y}' \in \mathcal{Y}$. Then, for any $\boldsymbol{x} \in \mathcal{X}$,

$$\begin{aligned}
\boldsymbol{x}^\top \boldsymbol{G}\boldsymbol{y}' &= \sum_{i \notin \text{supp}(\boldsymbol{x}^*)} x_i(\boldsymbol{G}\boldsymbol{y}')_i + \sum_{i \in \text{supp}(\boldsymbol{x}^*)} x_i(\boldsymbol{G}\boldsymbol{y}')_i \\
&\geq \sum_{i \notin \text{supp}(\boldsymbol{x}^*)} \left(x_i(\boldsymbol{G}\boldsymbol{y}^*)_i - x_i\|\boldsymbol{y}' - \boldsymbol{y}^*\|_1\right) + \sum_{i \in \text{supp}(\boldsymbol{x}^*)} \left(\frac{\xi}{2} \cdot x_i(\boldsymbol{G}(\boldsymbol{y} - \boldsymbol{y}^*))_i + x_i(\boldsymbol{G}\boldsymbol{y}^*)_i\right) \\
&\qquad \text{(using } G_{ij} \in [-1, 1] \text{ for the first part and } \boldsymbol{y}' = \boldsymbol{y}^* + \tfrac{\xi}{2N}(\boldsymbol{y} - \boldsymbol{y}^*) \text{ for the second)} \\
&\geq \sum_{i \notin \text{supp}(\boldsymbol{x}^*)} \left(x_i(\boldsymbol{G}\boldsymbol{y}^*)_i - x_i\|\boldsymbol{y}' - \boldsymbol{y}^*\|_1\right) + \sum_{i \in \text{supp}(\boldsymbol{x}^*)} x_i \left(V^*_{h(i)} - \sum_{g \in \mathcal{H}_i} V^*_g\right),
\end{aligned}$$

where the last inequality is due to Eq. (35) and Lemma 13. We continue to bound the terms above, which are bounded by

$$\begin{aligned}
&\geq \sum_{i \notin \text{supp}(\boldsymbol{x}^*)} \left(x_i\left((\boldsymbol{G}\boldsymbol{y}^*)_i - \xi\right)\right) + \sum_{i \in \text{supp}(\boldsymbol{x}^*)} x_i \left(V^*_{h(i)} - \sum_{g \in \mathcal{H}_i} V^*_g\right) \\
&\qquad \text{(using } \boldsymbol{y}' - \boldsymbol{y}^* = \tfrac{\xi}{2N}(\boldsymbol{y} - \boldsymbol{y}^*) \text{ and } \|\boldsymbol{y} - \boldsymbol{y}^*\|_1 \leq 2N) \\
&\geq \sum_{i \notin \text{supp}(\boldsymbol{x}^*)} x_i \left(V^*_{h(i)} - \sum_{g \in \mathcal{H}_i} V^*_g\right) + \sum_{i \in \text{supp}(\boldsymbol{x}^*)} x_i \left(V^*_{h(i)} - \sum_{g \in \mathcal{H}_i} V^*_g\right) \quad \text{(by the definition of } \xi) \\
&= \sum_i x_i \left(V^*_{h(i)} - \sum_{g \in \mathcal{H}_i} V^*_g\right).
\end{aligned}$$

The last term can be writen as the matrix form $\boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{V}^* = \mathbf{e}^\top \boldsymbol{V}^* = V_0^*$. This shows that $\min_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{x}^\top \boldsymbol{G}\boldsymbol{y}' \ge V_0^*$, that is, $\boldsymbol{y}' \neq \boldsymbol{y}^*$ is also a maximin point, contradicting the uniqueness assumption. Therefore, $c_y > 0$ has to hold, and so does $c_x > 0$ by the same argument. □

We continue to define more constants in the following.

Definition 4. *Define constants* $z_{\min} \triangleq \min_i \widehat{z}_{1,i} \in (0,1]$,

$$\epsilon_{van} \triangleq \min_{j\in supp(\boldsymbol{z}^*)} \exp\left(-\frac{P(1+\ln(1/z_{\min}))}{z_j^*}\right) \in (0,1),$$

$$\epsilon_{dil} \triangleq \min_{j\in supp(\boldsymbol{z}^*)} \left\{ z_j^* \cdot \exp\left(-\frac{\|\alpha\|_\infty P^2 \ln\left(1/z_{\min}\right)}{z_j^*}\right) \cdot \left(\frac{3}{4}\right)^P \right\} \in (0,1).$$

For all $i \in \mathrm{supp}(\boldsymbol{z}^*)$, we will show that ϵ_{van} is a lower bound of $\widehat{z}_{t,i}$ for VOMWU, while ϵ_{dil} is a lower bound of $\widehat{z}_{t,i}$ and $z_{t,i}$ for DOMWU. We show the results in Lemma 20 and defer the proof there.

E.1.3 Proof of Lemma 12

We are almost ready to prove Lemma 12. Before that, we first show the following auxiliary lemma.

Lemma 15. *For any* $\boldsymbol{z} \in \mathcal{Z}$*, we have*

$$\max_{\boldsymbol{z}'\in\mathcal{V}^*(\mathcal{Z})} F(\boldsymbol{z})^\top(\boldsymbol{z}-\boldsymbol{z}') \ge C\|\boldsymbol{z}^* - \boldsymbol{z}\|_1,$$

for $C = \min\{c_x, c_y\} \in (0,1]$.

Proof. Recall that $V_0^* = \boldsymbol{x}^{*\top}\boldsymbol{G}\boldsymbol{y}^*$ is the game value and note that

$$\begin{aligned}
\max_{\boldsymbol{z}'\in\mathcal{V}^*(\mathcal{Z})} F(\boldsymbol{z})^\top(\boldsymbol{z}-\boldsymbol{z}') &= \max_{\boldsymbol{z}'\in\mathcal{V}^*(\mathcal{Z})} (\boldsymbol{x}-\boldsymbol{x}')^\top \boldsymbol{G}\boldsymbol{y} + \boldsymbol{x}^\top\boldsymbol{G}(\boldsymbol{y}'-\boldsymbol{y}) = \max_{\boldsymbol{z}'\in\mathcal{V}^*(\mathcal{Z})} -\boldsymbol{x}'^\top\boldsymbol{G}\boldsymbol{y} + \boldsymbol{x}^\top\boldsymbol{G}\boldsymbol{y}' \\
&= \max_{\boldsymbol{x}'\in\mathcal{V}^*(\mathcal{X})} \left(V_0^* - \boldsymbol{x}'^\top\boldsymbol{G}\boldsymbol{y}\right) + \max_{\boldsymbol{y}'\in\mathcal{V}^*(\mathcal{Y})} \left(\boldsymbol{x}^\top\boldsymbol{G}\boldsymbol{y}' - V_0^*\right) \\
&= \max_{\boldsymbol{x}'\in\mathcal{V}^*(\mathcal{X})} \boldsymbol{x}'^\top\boldsymbol{G}(\boldsymbol{y}^*-\boldsymbol{y}) + \max_{\boldsymbol{y}'\in\mathcal{V}^*(\mathcal{Y})} (\boldsymbol{x}-\boldsymbol{x}^*)^\top\boldsymbol{G}\boldsymbol{y}' \\
&\ge c_y\|\boldsymbol{y}^*-\boldsymbol{y}\|_1 + c_x\|\boldsymbol{x}^*-\boldsymbol{x}\|_1 && \text{(by Definition 3)} \\
&\ge \min\{c_x,c_y\}\|\boldsymbol{z}^*-\boldsymbol{z}\|_1,
\end{aligned}$$

where the third equality is due to Eq. (29), $\boldsymbol{x}' \in \mathcal{V}^*(\mathcal{X})$, and

$$V_0^* = \boldsymbol{e}^\top\boldsymbol{V}^* = (\boldsymbol{x}'^\top\boldsymbol{A}^\top)\boldsymbol{V}^* = \boldsymbol{x}'^\top(\boldsymbol{A}^\top\boldsymbol{V}^*) = \boldsymbol{x}'^\top(\boldsymbol{G}\boldsymbol{y}^*);$$

the case for $\boldsymbol{y}'$ is similar. This completes the proof. □

Now we show the proof of Lemma 12.

Proof of Lemma 12. Below we consider any $\boldsymbol{z}' \in \mathcal{Z}$ such that $\mathrm{supp}(\boldsymbol{z}') \subseteq \mathrm{supp}(\boldsymbol{z}^*)$, that is, $\boldsymbol{z}' \in \mathcal{V}^*(\mathcal{Z})$. Considering Eq. (2) , and using the first-order optimality condition of $\widehat{\boldsymbol{z}}_{t+1}$, we have

$$(\nabla\psi(\widehat{\boldsymbol{z}}_{t+1}) - \nabla\psi(\widehat{\boldsymbol{z}}_t) + \eta F(\boldsymbol{z}_t))^\top(\boldsymbol{z}' - \widehat{\boldsymbol{z}}_{t+1}) \ge 0. \tag{36}$$

Rearranging the terms and we get

$$\eta F\left(\boldsymbol{z}_t\right)^\top\left(\widehat{\boldsymbol{z}}_{t+1} - \boldsymbol{z}'\right) \le \left(\nabla\psi(\widehat{\boldsymbol{z}}_{t+1}) - \nabla\psi(\widehat{\boldsymbol{z}}_t)\right)^\top\left(\boldsymbol{z}' - \widehat{\boldsymbol{z}}_{t+1}\right). \tag{37}$$

The left hand side of Eq. (37) is lower bounded as

$$\begin{aligned}
\eta F\left(\boldsymbol{z}_t\right)^\top\left(\widehat{\boldsymbol{z}}_{t+1} - \boldsymbol{z}'\right) &= \eta F\left(\widehat{\boldsymbol{z}}_{t+1}\right)^\top\left(\widehat{\boldsymbol{z}}_{t+1} - \boldsymbol{z}'\right) + \eta\left(F\left(\boldsymbol{z}_t\right) - F\left(\widehat{\boldsymbol{z}}_{t+1}\right)\right)^\top\left(\widehat{\boldsymbol{z}}_{t+1} - \boldsymbol{z}'\right) \\
&\ge \eta F\left(\widehat{\boldsymbol{z}}_{t+1}\right)^\top\left(\widehat{\boldsymbol{z}}_{t+1} - \boldsymbol{z}'\right) - \eta\|F\left(\boldsymbol{z}_t\right) - F\left(\widehat{\boldsymbol{z}}_{t+1}\right)\|_\infty\|\widehat{\boldsymbol{z}}_{t+1} - \boldsymbol{z}'\|_1 \\
&\ge \eta F\left(\widehat{\boldsymbol{z}}_{t+1}\right)^\top\left(\widehat{\boldsymbol{z}}_{t+1} - \boldsymbol{z}'\right) - 2P\eta\left\|\boldsymbol{z}_t - \widehat{\boldsymbol{z}}_{t+1}\right\|_1 \\
&\ge \eta F\left(\widehat{\boldsymbol{z}}_{t+1}\right)^\top\left(\widehat{\boldsymbol{z}}_{t+1} - \boldsymbol{z}'\right) - \frac{1}{4}\left\|\boldsymbol{z}_t - \widehat{\boldsymbol{z}}_{t+1}\right\|_1; && (\eta \le 1/(8P))
\end{aligned}$$

When $\psi = \Psi^{\text{van}}$, we have

$$(\nabla\Psi^{\text{van}}(\widehat{\boldsymbol{z}}_{t+1}) - \nabla\Psi^{\text{van}}(\widehat{\boldsymbol{z}}_t))_i = (1 + \ln \widehat{z}_{t+1,i}) - (1 + \ln \widehat{z}_{t,i}) = \ln \frac{\widehat{z}_{t+1,i}}{\widehat{z}_{t,i}}. \tag{38}$$

On the other hand, when $\psi = \Psi^{\text{dil}}_{\boldsymbol{\alpha}}$, by Eq. (23), we have

$$\left(\nabla\Psi^{\text{dil}}_{\boldsymbol{\alpha}}(\widehat{\boldsymbol{z}}_{t+1}) - \nabla\Psi^{\text{dil}}_{\boldsymbol{\alpha}}(\widehat{\boldsymbol{z}}_t)\right)_i = \alpha_i \ln \frac{\widehat{q}_{t+1,i}}{\widehat{q}_{t,i}}. \tag{39}$$

Therefore, the right hand side of Eq. (37) for VOMWU is upper bounded by

$$\begin{aligned}
&\left(\nabla\Psi^{\text{van}}(\widehat{\boldsymbol{z}}_{t+1}) - \nabla\Psi^{\text{van}}(\widehat{\boldsymbol{z}}_t)\right)^\top (\boldsymbol{z}' - \widehat{\boldsymbol{z}}_{t+1}) \\
&= \sum_i (z'_i - \widehat{z}_{t+1,i}) \ln \frac{\widehat{z}_{t+1,i}}{\widehat{z}_{t,i}} && \text{(Eq. (38))} \\
&\le \|\widehat{\boldsymbol{z}}_{t+1} - \widehat{\boldsymbol{z}}_t\|_1 - D_{\Psi^{\text{van}}}(\widehat{\boldsymbol{z}}_{t+1}, \widehat{\boldsymbol{z}}_t) + \sum_{i \in \text{supp}(\boldsymbol{z}^*)} z'_i \ln \frac{\widehat{z}_{t+1,i}}{\widehat{z}_{t,i}} && (\text{supp}(\boldsymbol{z}') \subseteq \text{supp}(\boldsymbol{z}^*)) \\
&\le \|\widehat{\boldsymbol{z}}_{t+1} - \widehat{\boldsymbol{z}}_t\|_1 + \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \left|\ln \frac{\widehat{z}_{t+1,i}}{\widehat{z}_{t,i}}\right| \\
&\le \|\widehat{\boldsymbol{z}}_{t+1} - \widehat{\boldsymbol{z}}_t\|_1 + \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \ln\left(1 + \frac{|\widehat{z}_{t+1,i} - \widehat{z}_{t,i}|}{\min\{\widehat{z}_{t+1,i}, \widehat{z}_{t,i}\}}\right) \\
&\le \|\widehat{\boldsymbol{z}}_{t+1} - \widehat{\boldsymbol{z}}_t\|_1 + \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \frac{|\widehat{z}_{t+1,i} - \widehat{z}_{t,i}|}{\min\{\widehat{z}_{t+1,i}, \widehat{z}_{t,i}\}} && (\ln(1+a) \le a) \\
&\le \frac{2}{\epsilon_{\text{van}}} \|\widehat{\boldsymbol{z}}_{t+1} - \widehat{\boldsymbol{z}}_t\|_1; && (\text{Lemma 20 and } \epsilon_{\text{van}} \le 1)
\end{aligned}$$

on the other hand, the right hand side of Eq. (37) for DOMWU is upper bounded by

$$\begin{aligned}
&\left(\nabla\Psi^{\text{dil}}_{\boldsymbol{\alpha}}(\widehat{\boldsymbol{z}}_{t+1}) - \nabla\Psi^{\text{dil}}_{\boldsymbol{\alpha}}(\widehat{\boldsymbol{z}}_t)\right)^\top (\boldsymbol{z}' - \widehat{\boldsymbol{z}}_{t+1}) \\
&= \sum_i \alpha_i (z'_i - \widehat{z}_{t+1,i}) \ln \frac{\widehat{q}_{t+1,i}}{\widehat{q}_{t,i}} && \text{(Eq. (39))} \\
&= \sum_i \alpha_i z'_i \ln \frac{\widehat{q}_{t+1,i}}{\widehat{q}_{t,i}} - \sum_i \alpha_i \widehat{z}_{t+1,i} \ln \frac{\widehat{q}_{t+1,i}}{\widehat{q}_{t,i}} \\
&= \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \alpha_i z'_i \ln \frac{\widehat{q}_{t+1,i}}{\widehat{q}_{t,i}} - \sum_{h \in \mathcal{H}^{\mathcal{Z}}} \alpha_h \widehat{z}_{t+1,\sigma(h)} \sum_{j \in \Omega_h} \widehat{q}_{t+1,j} \ln \frac{\widehat{q}_{t+1,j}}{\widehat{q}_{t,j}} \\
&\le \|\boldsymbol{\alpha}\|_\infty \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \left|\ln \frac{\widehat{q}_{t+1,i}}{\widehat{q}_{t,i}}\right| && \left(\textstyle\sum_{j \in \Omega_h} \widehat{q}_{t+1,j} \ln \frac{\widehat{q}_{t+1,j}}{\widehat{q}_{t,j}} \ge 0\right) \\
&\le \|\boldsymbol{\alpha}\|_\infty \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \ln\left(1 + \frac{|\widehat{q}_{t+1,i} - \widehat{q}_{t,i}|}{\min\{\widehat{q}_{t+1,i}, \widehat{q}_{t,i}\}}\right) \\
&\le \|\boldsymbol{\alpha}\|_\infty \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \frac{|\widehat{q}_{t+1,i} - \widehat{q}_{t,i}|}{\min\{\widehat{q}_{t+1,i}, \widehat{q}_{t,i}\}} && (\ln(1+a) \le a) \\
&\le \frac{\|\boldsymbol{\alpha}\|_\infty}{\epsilon_{\text{dil}}} \sum_{i \in \text{supp}(\boldsymbol{z}^*)} |\widehat{q}_{t+1,i} - \widehat{q}_{t,i}|. && \text{(Lemma 20)}
\end{aligned}$$

Since $\boldsymbol{z}'$ can be chosen as any point in $\mathcal{V}^*(\mathcal{Z})$, we further lower bound the left-hand side of Eq. (37) using Lemma 15 and get for VOMWU,

$$\begin{aligned}
\eta C \|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_{t+1}\|_1 &\le \frac{2}{\epsilon_{\text{van}}} \|\widehat{\boldsymbol{z}}_{t+1} - \widehat{\boldsymbol{z}}_t\|_1 + \frac{\|\boldsymbol{z}_t - \widehat{\boldsymbol{z}}_{t+1}\|_1}{4} \\
&\le \frac{2}{\epsilon_{\text{van}}} \left(\|\widehat{\boldsymbol{z}}_{t+1} - \widehat{\boldsymbol{z}}_t\|_1 + \|\boldsymbol{z}_t - \widehat{\boldsymbol{z}}_{t+1}\|_1\right),
\end{aligned} \tag{40}$$

and for DOMWU,

$$\eta C\|\boldsymbol{z}^*-\widehat{\boldsymbol{z}}_{t+1}\|_1 \le \frac{1}{4}\|\boldsymbol{z}_t-\widehat{\boldsymbol{z}}_{t+1}\|_1 + \frac{\|\boldsymbol{\alpha}\|_\infty}{\epsilon_{\text{dil}}}\sum_{i\in\text{supp}(\boldsymbol{z}^*)}|\widehat{q}_{t+1,i}-\widehat{q}_{t,i}|, \qquad \text{(Lemma 20)}$$

$$\le \frac{1}{4}\|\boldsymbol{z}_t-\widehat{\boldsymbol{z}}_{t+1}\|_1 + \frac{\|\boldsymbol{\alpha}\|_\infty}{\epsilon_{\text{dil}}}\sum_{i\in\text{supp}(\boldsymbol{z}^*)}\frac{|\widehat{z}_{t+1,i}-\widehat{z}_{t,i}|}{\widehat{z}_{t+1,p_i}} + \widehat{z}_{t,i}\frac{|\widehat{z}_{t+1,p_i}-\widehat{z}_{t,p_i}|}{\widehat{z}_{t+1,p_i}\widehat{z}_{t,p_i}}$$

$$\le \frac{1}{4}\|\boldsymbol{z}_t-\widehat{\boldsymbol{z}}_{t+1}\|_1 + \frac{\|\boldsymbol{\alpha}\|_\infty}{\epsilon_{\text{dil}}^2}\sum_{i\in\text{supp}(\boldsymbol{z}^*)}|\widehat{z}_{t+1,i}-\widehat{z}_{t,i}| + |\widehat{z}_{t+1,p_i}-\widehat{z}_{t,p_i}| \qquad \text{(Lemma 20)}$$

$$\le \frac{1}{4}\|\boldsymbol{z}_t-\widehat{\boldsymbol{z}}_{t+1}\|_1 + \frac{P\|\boldsymbol{\alpha}\|_\infty}{\epsilon_{\text{dil}}^2}\|\widehat{\boldsymbol{z}}_{t+1}-\widehat{\boldsymbol{z}}_t\|_1$$

$$\le \left(\frac{1}{4}+\frac{P\|\boldsymbol{\alpha}\|_\infty}{\epsilon_{\text{dil}}^2}\right)\left(\|\widehat{\boldsymbol{z}}_{t+1}-\widehat{\boldsymbol{z}}_t\|_1 + \|\boldsymbol{z}_t-\widehat{\boldsymbol{z}}_{t+1}\|_1\right). \qquad (41)$$

Squaring both sides of Eq. (40), we get

$$\eta^2C^2\|\boldsymbol{z}^*-\widehat{\boldsymbol{z}}_{t+1}\|_1^2 \le \frac{4}{\epsilon_{\text{van}}^2}\left(\|\widehat{\boldsymbol{z}}_{t+1}-\widehat{\boldsymbol{z}}_t\|_1 + \|\boldsymbol{z}_t-\widehat{\boldsymbol{z}}_{t+1}\|_1\right)^2$$

$$\le \frac{8}{\epsilon_{\text{van}}^2}\left(\|\widehat{\boldsymbol{z}}_{t+1}-\widehat{\boldsymbol{z}}_t\|_1^2 + \|\boldsymbol{z}_t-\widehat{\boldsymbol{z}}_{t+1}\|_1^2\right). \qquad (42)$$

Using the strong convexity of the regularizers, the left hand side of Eq. (25) can be bounded by

$$D_\psi(\widehat{\boldsymbol{z}}_{t+1},\boldsymbol{z}_t) + D_\psi(\boldsymbol{z}_t,\widehat{\boldsymbol{z}}_t)$$

$$\ge \frac{1}{2}\|\widehat{\boldsymbol{z}}_{t+1}-\boldsymbol{z}_t\|_1^2 + \frac{1}{2}\|\boldsymbol{z}_t-\widehat{\boldsymbol{z}}_t\|_1^2 \qquad (a^2+b^2\ge\tfrac{1}{2}(a+b)^2)$$

$$\ge \frac{1}{8}\|\widehat{\boldsymbol{z}}_{t+1}-\boldsymbol{z}_t\|_1^2 + \frac{1}{4}\left(\|\widehat{\boldsymbol{z}}_{t+1}-\boldsymbol{z}_t\|_1^2 + \|\boldsymbol{z}_t-\widehat{\boldsymbol{z}}_t\|_1^2\right)$$

$$\ge \frac{1}{8}\left(\|\widehat{\boldsymbol{z}}_{t+1}-\widehat{\boldsymbol{z}}_t\|_1^2 + \|\widehat{\boldsymbol{z}}_{t+1}-\boldsymbol{z}_t\|_1^2\right) \qquad (a^2+b^2\ge\tfrac{1}{2}(a+b)^2 \text{ and triangle inequality})$$

Combining this with Eq. (41) finishes the proof for VOMWU. The similar argument works for DOMWU by combining the inequality above with Eq. (42). □

E.2 Proofs of Eq. (12) and Eq. (13)

In this subsection, we prove Eq. (12) and Eq. (13). We first show Lemma 16 and Lemma 17. Then we get Eq. (12) and Eq. (13) by substituting $\boldsymbol{z}$ with $\widehat{\boldsymbol{z}}_{t+1}$ in these lemmas.

Lemma 16. *For any $\boldsymbol{z}\in\mathcal{Z}$, we have*

$$D_{\Psi^{van}}(\boldsymbol{z}^*,\boldsymbol{z}) \le \sum_{i\in supp(\boldsymbol{z}^*)}\frac{(z_i^*-z_i)^2}{\min_{i\in supp(\boldsymbol{z}^*)} z_i} + \sum_{i\notin supp(\boldsymbol{z}^*)} z_i \le \frac{3P\|\boldsymbol{z}^*-\boldsymbol{z}\|}{\min_{i\in supp(\boldsymbol{z}^*)} z_i}.$$

Proof. By Eq. (14), we have

$$D_{\Psi^{\text{van}}}(\boldsymbol{z}^*,\boldsymbol{z}) \le \sum_i \frac{(z_i^*-z_i)^2}{z_i} \le \sum_{i\in\text{supp}(\boldsymbol{z}^*)}\frac{(z_i^*-z_i)^2}{\min_{i\in\text{supp}(\boldsymbol{z}^*)} z_i} + \sum_{i\notin\text{supp}(\boldsymbol{z}^*)} z_i$$

$$\le \frac{\|\boldsymbol{z}^*-\boldsymbol{z}\|^2}{\min_{i\in\text{supp}(\boldsymbol{z}^*)} z_i} + \|\boldsymbol{z}^*-\boldsymbol{z}\|_1 \le \frac{3P\|\boldsymbol{z}^*-\boldsymbol{z}\|}{\min_{i\in\text{supp}(\boldsymbol{z}^*)} z_i},$$

where the last inequality is because $\|\boldsymbol{z}^*-\boldsymbol{z}\|\le 2P$ and $\|\boldsymbol{z}^*-\boldsymbol{z}\|_1\le P\|\boldsymbol{z}^*-\boldsymbol{z}\|$. □

Lemma 17. *For any $\boldsymbol{z}\in\mathcal{Z}$, we have*

$$D_{\Psi^{dil}_{\boldsymbol{\alpha}}}(\boldsymbol{z}^*,\boldsymbol{z}) \le \|\boldsymbol{\alpha}\|_\infty\left(\sum_{i\in supp(\boldsymbol{z}^*)}\frac{4P}{z_i^*}\frac{(z_i^*-z_i)^2}{q_i} + \sum_{i\notin supp(\boldsymbol{z}^*)} z_{p_i}^* q_i\right) \le \frac{4P\|\boldsymbol{\alpha}\|_\infty}{\min_{i\in supp(\boldsymbol{z}^*)} z_i^* z_i}\|\boldsymbol{z}^*-\boldsymbol{z}\|_1.$$

Proof. By direction calculation and [Wei et al., 2021, Lemma 16], we have

$$
\begin{aligned}
D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}^*, \boldsymbol{z}) &= \sum_i \alpha_{h(i)} z^*_{p_i} q^*_i \ln\left(\frac{q^*_i}{q_i}\right) && \text{(Eq. (24))} \\
&\leq \|\boldsymbol{\alpha}\|_\infty \sum_i z^*_{p_i} \left(\mathbb{1}\{i \in \text{supp}(\boldsymbol{z}^*)\} \frac{(q^*_i - q_i)^2}{q_i} + \mathbb{1}\{i \notin \text{supp}(\boldsymbol{z}^*)\} q_i \right) \\
&\leq \|\boldsymbol{\alpha}\|_\infty \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \frac{2 z^*_{p_i}}{q_i} \left(\left(\frac{z^*_i - z_i}{z^*_{p_i}} \right)^2 + \left(\frac{z_i}{z^*_{p_i}} - \frac{z_i}{z_{p_i}} \right)^2 \right) + \|\boldsymbol{\alpha}\|_\infty \sum_{i \notin \text{supp}(\boldsymbol{z}^*)} z^*_{p_i} q_i \\
&= \|\boldsymbol{\alpha}\|_\infty \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \left(\frac{2}{q_i z^*_{p_i}} (z^*_i - z_i)^2 + \frac{2 q_i}{z^*_{p_i}} (z^*_{p_i} - z_{p_i})^2 \right) + \|\boldsymbol{\alpha}\|_\infty \sum_{i \notin \text{supp}(\boldsymbol{z}^*)} z^*_{p_i} q_i \\
&\leq \|\boldsymbol{\alpha}\|_\infty \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \left(\frac{2}{z^*_{p_i}} \frac{(z^*_i - z_i)^2}{q_i} + \frac{2P}{z^*_i} (z^*_i - z_i)^2 \right) + \|\boldsymbol{\alpha}\|_\infty \sum_{i \notin \text{supp}(\boldsymbol{z}^*)} z^*_{p_i} q_i \\
& && (p_i \in \text{supp}(\boldsymbol{z}^*) \text{ for all } i \in \text{supp}(\boldsymbol{z}^*)) \\
&\leq \|\boldsymbol{\alpha}\|_\infty \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \left(\frac{4P}{z^*_i} \frac{(z^*_i - z_i)^2}{q_i} \right) + \|\boldsymbol{\alpha}\|_\infty \sum_{i \notin \text{supp}(\boldsymbol{z}^*)} z^*_{p_i} q_i.
\end{aligned}
$$

This proves the first inequality. The second equality in the lemma follows from

$$
\begin{aligned}
D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}^*, \boldsymbol{z}) &\leq \|\boldsymbol{\alpha}\|_\infty \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \left(\frac{4P}{z^*_i} \frac{(z^*_i - z_i)^2}{q_i} \right) + \|\boldsymbol{\alpha}\|_\infty \sum_{i \notin \text{supp}(\boldsymbol{z}^*)} z^*_{p_i} q_i \\
&\leq \|\boldsymbol{\alpha}\|_\infty \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \left(\frac{4P}{z^*_i z_i} \right) |z^*_i - z_i| + \|\boldsymbol{\alpha}\|_\infty \sum_{i \notin \text{supp}(\boldsymbol{z}^*)} \frac{z^*_{p_i} z_i}{z_{p_i}} && (|z^*_i - z_i| \leq 1) \\
&\leq \|\boldsymbol{\alpha}\|_\infty \sum_{i \in \text{supp}(\boldsymbol{z}^*)} \left(\frac{4P}{z^*_i z_i} \right) |z^*_i - z_i| + \|\boldsymbol{\alpha}\|_\infty \sum_{i \notin \text{supp}(\boldsymbol{z}^*)} \frac{\mathbb{1}\{p_i \in \text{supp}(\boldsymbol{z}^*)\}}{z_{p_i}} \cdot |z_i - 0| \\
&\leq \frac{4P\|\boldsymbol{\alpha}\|_\infty}{\min_{i \in \text{supp}(\boldsymbol{z}^*)} z^*_i z_i} \|\boldsymbol{z}^* - \boldsymbol{z}\|_1.
\end{aligned}
$$

□

E.3 Lower Bounds on the Probability Masses

In this subsection, we show for all $i \in \text{supp}(\boldsymbol{z}^*)$ and t, $\widehat{z}_{t,i}$ computed by VOMWU can be lower bounded by ϵ_{van}, while $\widehat{z}_{t,i}$ and $z_{t,i}$ computed by DOMWU can be lower bounded by ϵ_{dil}, where ϵ_{van} and ϵ_{dil} are defined in Definition 4. We state the results in Lemma 19 and Lemma 20, respectively. We first state the stability of $\widehat{\boldsymbol{q}}_t$ and $\boldsymbol{q}_t$, which directly follows from the stability of OMWU on simplex, for example, [Wei et al., 2021, Lemma 17].

Lemma 18. *For $\eta \leq \frac{1}{8P}$,* DOMWU *guarantees $\frac{3}{4}\widehat{q}_{t,i} \leq q_{t,i} \leq \frac{4}{3}\widehat{q}_{t,i}$ and $\frac{3}{4}\widehat{q}_{t,i} \leq \widehat{q}_{t+1,i} \leq \frac{4}{3}\widehat{q}_{t,i}$.*

Lemma 19. *For all $i \in \text{supp}(\boldsymbol{z}^*)$ and t,* VOMWU *guarantees that $\widehat{z}_{t,i} \geq \epsilon_{van}$.*

Proof. Using Eq. (5), we have

$$
D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_t) \leq \Theta_t \leq \cdots \leq \Theta_1 = \tfrac{1}{16} D_\psi(\widehat{\boldsymbol{z}}_1, \boldsymbol{z}_0) + D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1) = D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1), \tag{43}
$$

where the last equality is because $\widehat{\boldsymbol{z}}_1 = \boldsymbol{z}_0$. Thus, $D_{\Psi^{\text{van}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_t) \le D_{\Psi^{\text{van}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1)$. Then, for any $i \in \text{supp}(\boldsymbol{z}^*)$, we have

$$\begin{aligned}
z_i^* \ln \frac{1}{\widehat{z}_{t,i}} &\le \sum_j z_j^* \ln \frac{1}{\widehat{z}_{t,j}} = D_{\Psi^{\text{van}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_t) + \sum_j \left(z_j^* - \widehat{z}_{t,j} - z_j^* \ln z_j^*\right) \\
&\le D_{\Psi^{\text{van}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1) - \sum_j z_j^* \ln z_j^* + \sum_j (z_j^* - \widehat{z}_{t,j}) \\
&\le \sum_j z_j^* \ln \frac{1}{\widehat{z}_{1,j}} + \sum_j (\widehat{z}_{1,j} - \widehat{z}_{t,j}) \\
&\le P + \sum_j z_j^* \ln \frac{1}{\widehat{z}_{1,j}} = P(1 + \ln(1/z_{\min})).
\end{aligned}$$

Therefore, we conclude for all t and $i \in \text{supp}(\boldsymbol{z}^*)$, $\widehat{z}_{t,i}$ satisfies

$$\widehat{z}_{t,i} \ge \exp\left(-\frac{P(1+\ln(1/z_{\min}))}{z_i^*}\right) \ge \min_{j \in \text{supp}(\boldsymbol{z}^*)} \exp\left(-\frac{P(1+\ln(1/z_{\min}))}{z_j^*}\right) = \epsilon_{\text{van}}.$$

□

Lemma 20. *For all $i \in supp(\boldsymbol{z}^*)$ and t,* DOMWU *guarantees that $\widehat{q}_{t,i} \ge \widehat{z}_{t,i} \ge \epsilon_{dil}$ and $q_{t,i} \ge z_{t,i} \ge \epsilon_{dil}$.*

Proof. Similar to Lemma 19, applying Eq. (43) gives $D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_t) \le D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1)$. Then, for any $i \in \text{supp}(\boldsymbol{z}^*)$, we have

$$\begin{aligned}
z_i^* \ln \frac{1}{\widehat{q}_{t,i}} &\le \sum_j \alpha_j z_j^* \ln \frac{1}{\widehat{q}_{t,j}} = D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_t) - \sum_j \alpha_j z_j^* \ln q_j^* \le D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1) - \sum_j \alpha_j z_j^* \ln q_j^* \\
&= \sum_j \alpha_j z_j^* \ln \frac{1}{\widehat{q}_{1,j}} \le \|\boldsymbol{\alpha}\|_\infty P \ln(1/z_{\min}).
\end{aligned}$$

Therefore, we conclude for all t and $i \in \text{supp}(\boldsymbol{z}^*)$, $\widehat{q}_{t,i}$ satisfies

$$\widehat{q}_{t,i} \ge \exp\left(-\frac{\|\boldsymbol{\alpha}\|_\infty P \ln(1/z_{\min})}{z_i^*}\right) \ge \min_{j \in \text{supp}(\boldsymbol{z}^*)} \exp\left(-\frac{\|\boldsymbol{\alpha}\|_\infty P \ln(1/z_{\min})}{z_j^*}\right)$$

and $\widehat{z}_{t,i}$ satisfies

$$\widehat{z}_{t,i} = \widehat{z}_{t,p_i}\widehat{q}_{t,i} = \widehat{z}_{t,p_{p_i}}\widehat{q}_{t,p_i}\widehat{q}_{t,i} \ge \prod_{j \in \text{supp}(\boldsymbol{z}^*)} \widehat{q}_{t,j} \ge \min_{j \in \text{supp}(\boldsymbol{z}^*)} \exp\left(-\frac{\|\alpha\|_\infty P \ln(1/z_{\min})}{z_j^*}\right)^P.$$

This finishes the first part of the proof. Finally, using Lemma 18, we have for $i \in \text{supp}(\boldsymbol{z}^*)$, $q_{t,i} \ge \frac{3}{4}\widehat{q}_{t,i} \ge \epsilon_{\text{dil}}$ and

$$z_{t,i} \ge \prod_{j \in \text{supp}(\boldsymbol{z}^*)} q_{t,j} \ge \min_{j \in \text{supp}(\boldsymbol{z}^*)} \exp\left(-\frac{\|\alpha\|_\infty P^2 \ln(1/z_{\min})}{z_j^*}\right) \cdot \left(\frac{3}{4}\right)^P \ge \epsilon_{\text{dil}}.$$

□

E.4 Proof of Theorem 6

Based on the results in the previous subsections and the discussion in Section 6.3, we can get $\Theta_t = \mathcal{O}(1/t)$ for both VOMWU and DOMWU. In this subsection, we formally state the results in Theorem 21 for both VOMWU and DOMWU, and show the proof by combining all the components. In particular, the result for VOMWU implies Theorem 6.

Theorem 21. *Under the uniqueness assumption,* VOMWU *and* DOMWU *with step size* $\eta \le \frac{1}{8P}$ *guarantee* $D_{\Psi^{\text{van}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_t) \le \frac{C_{13}}{t}$ *and* $D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_t) \le \frac{C'_{13}}{t}$*, respectively, where* $C'_{13}, C_{13} > 0$ *are some constants depending on the game,* $\widehat{\boldsymbol{z}}_1$*, and* η*.*

Proof. We start from Lemma 12. Using Lemma 16 and Lemma 19, the right hand side of Eq. (25) can be bounded by

$$\zeta_t \ge C_{12}\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_{t+1}\|^2 \ge \frac{\epsilon_{\text{van}}^2 C_{12}}{9P^2} D_{\Psi^{\text{van}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1})^2, \tag{44}$$

for VOMWU. Similarly, using Lemma 17 and Lemma 20, we have for DOMWU,

$$\zeta_t \ge C_{12}\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_{t+1}\|^2 \ge \frac{\epsilon_{\text{van}}^4 C_{12}}{16P^2\|\boldsymbol{\alpha}\|_\infty^2} D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1})^2.$$

On the other hand, applying Eq. (5) repeatedly, we get

$$D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1) = \Theta_1 \ge \cdots \ge \Theta_{t+1} \ge \frac{1}{16} D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t).$$

Thus, $\zeta_t \ge D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) \ge C_{10} D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t)^2$ for some $C_{10} > 0$ depending on $D_\psi(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1)$. Combining this with Eq. (44) gives

$$\begin{aligned}
\zeta_t &= \frac{1}{2}\left(D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) + D_\psi(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t)\right) + \frac{1}{2}\left(D_\psi(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) + D_\psi(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t)\right) \\
&\ge \frac{1}{2}\cdot\frac{\epsilon_{\text{van}}^2 C_{12}}{9P^2} D_{\Psi^{\text{van}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1})^2 + \frac{1}{2}\cdot C_{10} D_{\Psi^{\text{van}}}(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t)^2 && \text{(Eq. (44))} \\
&\ge \frac{\epsilon_{\text{van}}^2 C_{12}}{36P^2}\cdot 2D_{\Psi^{\text{van}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1})^2 + \frac{C_{10}}{4}\cdot 2D_{\Psi^{\text{van}}}(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t)^2 \\
&\ge \min\left\{\frac{\epsilon_{\text{van}}^2 C_{12}}{36P^2}, \frac{C_{10}}{4}\right\}\Theta_{t+1}^2 \ge C_{11}\Theta_{t+1}^2,
\end{aligned}$$

for some $C_{11} > 0$. Similarly, we have $\zeta_t \ge C'_{11}\Theta_{t+1}^2$ for DOMWU and constant $C'_{11} > 0$. Applying this to Eq. (5), we obtain the recursion $\Theta_{t+1} \le \Theta_t - \frac{15}{16}C_{11}\Theta_{t+1}^2$. This implies $\Theta_t \le \frac{C_{13}}{t}$ for some constant C_{13} by [Wei et al., 2021, Lemma 12]. With the same argument, we can prove the case for DOMWU. □

E.5 Proof of Theorem 7

E.5.1 The Significant Difference Lemma

In this subsection, we explain how to get the linear convergence result of DOMWU. As we discuss in the end of Section 6.3, this requires showing that $\sum_{i \notin \text{supp}(\boldsymbol{z}^*)} z^*_{p_i}\widehat{q}_{t+1,i}$ decreases significantly as $\widehat{\boldsymbol{z}}_t$ gets close enough to $\boldsymbol{z}^*$. The argument is shown in Lemma 24. Before that, we first show a lemma stating that for any information set $h \in \mathcal{H}^{\mathcal{Z}}$, indices $i, j \in \Omega_h$ such that $i \notin \text{supp}(\boldsymbol{z}^*)$ and $j \in \text{supp}(\boldsymbol{z}^*)$, $\widehat{L}_{t,i}$ is significantly larger than $\widehat{L}_{t,j}$ when $\widehat{\boldsymbol{z}}_t$ is close to $\boldsymbol{z}^*$.

Lemma 22. *Suppose* $\|\boldsymbol{z}^* - \boldsymbol{z}\| \le \frac{\eta^2\xi\epsilon_{\text{dil}}^2}{40P^3\|\boldsymbol{\alpha}\|_\infty}$*. Then for* $i \in \Omega_h$*,* $h \in \mathcal{H}^{\mathcal{X}}$*, we have*

$$\forall i \in \textit{supp}(\boldsymbol{x}^*), \qquad L_i \le V_h^* + \frac{\eta\xi}{10}; \qquad \forall i \notin \textit{supp}(\boldsymbol{x}^*), \qquad L_i \ge V_h^* + \frac{9\eta\xi}{10}, \tag{45}$$

where $L_i = (\boldsymbol{G}\boldsymbol{y})_i + \sum_{g \in \mathcal{H}_i} -\frac{\alpha_g}{\eta}\ln\left(\sum_{j\in\Omega_g} q_j \exp(-\eta L_j/\alpha_g)\right)$*,* $q_j = z_j/z_{p_j}$*.*

Proof. We first consider terminal index i. By the assumption $\|\boldsymbol{z}^* - \boldsymbol{z}\| \le \frac{\eta^2\xi\epsilon_{\text{dil}}^2}{40P^3\|\boldsymbol{\alpha}\|_\infty}$ and Lemma 13, we have $\|\boldsymbol{y} - \boldsymbol{y}^*\|_1 \le \frac{\eta\xi}{10P}$, and

$$L_i = (\boldsymbol{G}\boldsymbol{y})_i \le (\boldsymbol{G}\boldsymbol{y}^*)_i + \frac{\eta\xi}{10P} = V_h^* + \frac{\eta\xi}{10P} \tag{46}$$

for $i \in \text{supp}(\boldsymbol{x}^*)$ and

$$L_i = (\boldsymbol{G}\boldsymbol{y})_i \geq (\boldsymbol{G}\boldsymbol{y}^*)_i - \frac{\eta\xi}{10P} \geq V_h^* + \xi - \frac{\eta\xi}{10P} \geq V_h^* + \frac{9\eta\xi}{10}$$

for $i \notin \text{supp}(\boldsymbol{x}^*)$ by the definition of ξ. Therefore, this shows Eq. (45) for terminal indices. In the following, we complete the proof by backward induction. Specifically, for nonterminal index $i \notin \text{supp}(\boldsymbol{x}^*)$, we assume $L_j \geq V_{h(j)}^* + \frac{9\eta\xi}{10}$ for every *descendant* j (we say that index j is a descendant of index i if there exists a sequence of indexes $s_0, \ldots, s_K$ for some $K > 0$ such that $s_0 = j$, $s_K = i$, and $p_{s_{k-1}} = s_k$ for every $k \in [K]$). Note that we always have $j \notin \text{supp}(\boldsymbol{x}^*)$. We will prove L_i satisfies Eq. (45), which completes the proof for $i \notin \text{supp}(\boldsymbol{x}^*)$ by induction. By assumption, we have

$$\begin{aligned}
L_i &= (\boldsymbol{G}\boldsymbol{y})_i + \sum_{g \in \mathcal{H}_i} -\frac{\alpha_g}{\eta} \ln\left(\sum_{j \in \Omega_g} q_j \exp(-\eta L_j/\alpha_g)\right) \\
&\geq (\boldsymbol{G}\boldsymbol{y})_i + \sum_{g \in \mathcal{H}_i} -\frac{\alpha_g}{\eta} \ln\left(\max_{j \in \Omega_g} \exp(-\eta L_j/\alpha_g)\right) \\
&= (\boldsymbol{G}\boldsymbol{y})_i + \sum_{g \in \mathcal{H}_i} \min_{j \in \Omega_g} L_j \\
&\geq (\boldsymbol{G}\boldsymbol{y})_i + \sum_{g \in \mathcal{H}_i} V_g^* + \frac{9\eta\xi}{10} && \text{(by the induction hypothesis)} \\
&\geq (\boldsymbol{G}\boldsymbol{y}^*)_i - \xi + \sum_{g \in \mathcal{H}_i} V_g^* + \frac{9\eta\xi}{10} && (\|\boldsymbol{y} - \boldsymbol{y}^*\| \leq \xi) \\
&\geq V_{h(i)}^* + \frac{9\eta\xi}{10}. && \text{(Lemma 13)}
\end{aligned}$$

Similarly, for nonterminal index $i \in \text{supp}(\boldsymbol{x}^*)$, we show for every descendant j,

$$L_j \leq V_{h(j)}^* + \frac{\eta\xi}{10P} f(h(j)), \tag{47}$$

where $f : \mathcal{H}^{\mathcal{X}} \to \mathbb{R}^+$ is defined recursively as follows. For information set (simplex) g such that Ω_g contains terminal indices only, we let $f(g) = 1$. Otherwise, we define

$$f(g) = \max_{k \in \Omega_g} \sum_{s \in \mathcal{H}_k} \left(f(s) + \frac{1}{P}\right). \tag{48}$$

This shows Eq. (45) as Lemma 23 guarantees

$$f(h) \leq (P-1) \cdot \left(1 + \frac{1}{P}\right) < P,$$

for every simplex h. It remains to prove Eq. (48) by induction. For the base case that i is a terminal index, Eq. (47) clearly holds by Eq. (46). For nonterminal index i, we have

$$\begin{aligned}
L_i &= (\boldsymbol{G}\boldsymbol{y})_i + \sum_{g \in \mathcal{H}_i} -\frac{\alpha_g}{\eta} \ln\left(\sum_{j \in \Omega_g} q_j \exp(-\eta L_j/\alpha_g)\right) \\
&\leq (\boldsymbol{G}\boldsymbol{y})_i + \sum_{g \in \mathcal{H}_i} -\frac{\alpha_g}{\eta} \ln\left(\sum_{j \in \Omega_g \cap \text{supp}(\boldsymbol{x}^*)} q_j \exp(-\eta L_j/\alpha_g)\right) \\
&\leq (\boldsymbol{G}\boldsymbol{y})_i + \sum_{g \in \mathcal{H}_i} -\frac{\alpha_g}{\eta} \ln\left[\exp\left(-\eta\left(V_g^* + \frac{\eta\xi}{10P} f(g)\right)/\alpha_g\right) \sum_{j \in \Omega_g \cap \text{supp}(\boldsymbol{x}^*)} q_j\right] \\
& && \text{(by the assumption)} \\
&= (\boldsymbol{G}\boldsymbol{y})_i + \sum_{g \in \mathcal{H}_i} \left[V_g^* + \frac{\eta\xi}{20P} f(g) - \frac{\alpha_g}{\eta} \ln\left(\sum_{j \in \Omega_g \cap \text{supp}(\boldsymbol{x}^*)} q_j\right)\right]. && (49)
\end{aligned}$$

We continue to bound the last term. Let $c = \frac{\eta^2 \xi \epsilon_{\text{dil}}^2}{40P^3\|\boldsymbol{\alpha}\|_\infty}$. We have

$$
\begin{aligned}
-\frac{\alpha_g}{\eta} \ln\left(\sum_{j\in\Omega_g\cap\text{supp}(\boldsymbol{x}^*)} q_j\right) &= -\frac{\alpha_g}{\eta}\ln\left(\frac{\sum_{j\in\Omega_g\cap\text{supp}(\boldsymbol{x}^*)} x_j}{x_{p_j}}\right) \\
&\le -\frac{\alpha_g}{\eta}\ln\left(\frac{\sum_{j\in\Omega_g\cap\text{supp}(\boldsymbol{x}^*)}(x_j^* - c)}{x_{p_j}^* + c}\right) && (\|\boldsymbol{z}^* - \boldsymbol{z}\| \le \tfrac{\eta^2\xi\epsilon_{\text{dil}}^2}{40P^3\|\boldsymbol{\alpha}\|_\infty}) \\
&\le -\frac{\alpha_g}{\eta}\ln\left(\frac{x_{p_j}^* - c|\Omega_g|}{x_{p_j}^* + c}\right) \\
&= -\frac{\alpha_g}{\eta}\ln\left(1 - \frac{c(|\Omega_g|+1)}{x_{p_j}^* + c}\right) \\
&\le -\frac{\alpha_g}{\eta}\ln\left(1 - \frac{cP}{\epsilon_{\text{dil}}}\right) && (x_{p_j}^* + c \ge x_{p_j}^* \ge \epsilon_{\text{dil}} \text{ and } |\Omega_g| + 1 \le P) \\
&\le -\frac{\alpha_g}{\eta}\ln\left(1 - \frac{\eta^2\xi}{40\alpha_g P^2}\right) && (\text{by definition of } c) \\
&\le \frac{\eta\xi}{20P^2}. && (-\ln(1-x) < 2x \text{ for } 0 < x < 0.5)
\end{aligned}
$$

Plugging this back to the original inequalities, we get

$$
\begin{aligned}
L_i &\le (\boldsymbol{Gy})_i + \sum_{g\in\mathcal{H}_i}\left(V_g^* + \frac{\eta\xi}{10P}f(g) + \frac{\eta\xi}{20P^2}\right) \\
&\le (\boldsymbol{Gy}^*)_i + \frac{\eta\xi}{20P^2} + \sum_{g\in\mathcal{H}_i}\left(V_g^* + \frac{\eta\xi}{10P}f(g) + \frac{\eta\xi}{20P^2}\right) && (\|\boldsymbol{y} - \boldsymbol{y}^*\|_1 \le \tfrac{\eta\xi}{10P^2}) \\
&\le (\boldsymbol{Gy}^*)_i + \sum_{g\in\mathcal{H}_i}\left(V_g^* + \frac{\eta\xi}{10P}f(g) + \frac{\eta\xi}{10P^2}\right) && (i \text{ is nonterminal}) \\
&= (\boldsymbol{Gy}^*)_i + \sum_{g\in\mathcal{H}_i} V_g^* + \frac{\eta\xi}{10P}\sum_{g\in\mathcal{H}_i}\left(f(g) + \frac{1}{P}\right) \\
&\le V_h^* + \frac{\eta\xi}{10P}\sum_{g\in\mathcal{H}_i}\left(f(g) + \frac{1}{P}\right) && (\text{Lemma 13}) \\
&\le V_h^* + \frac{\eta\xi}{10P}f(h(i)), && (\text{Eq. (48)})
\end{aligned}
$$

which shows Eq. (47) by induction, and thus shows Eq. (45). □

Lemma 23. *Define $f : \mathcal{H}^{\mathcal{X}} \to \mathbb{R}^+$ as follows.*

$$
f(g) = \begin{cases} 1, & \text{if } \Omega_g \text{ contains terminal indices only;} \\ \max_{k\in\Omega_g} \sum_{s\in\mathcal{H}_k}\left(f(s) + \frac{1}{P}\right), & \text{otherwise.} \end{cases}
$$

Then for every $g \in \mathcal{H}^{\mathcal{X}}$, we have

$$
f(g) \le I_g\left(1 + \frac{1}{P}\right), \tag{50}
$$

where I_g is the number of indices that are the descendants of g (we say that index j is a descendant of simplex g if $j \in \Omega_g$ or index j is a descendant of index i for some $i \in \Omega_g$).

Proof. If Ω_g contains terminal indices only, since $I_g \ge 1$, Eq. (50) holds. Otherwise, suppose Eq. (50) holds for all simplexes that are descendants of g (we say that simplex h is a descendant of simplex g if there exists a sequence of simplexes $s_0, \ldots, s_K$ for some $K > 0$ such that $s_0 = h$, $s_K = g$, and $\sigma(s_{k-1}) \in \Omega_{s_k}$ for every $k \in [K]$). We define

$$k^* = \operatorname*{argmax}_{k \in \Omega_g} \sum_{s \in \mathcal{H}_k} \left(f(s) + \frac{1}{P} \right).$$

Then we have

$$\begin{aligned}
f(g) &= \sum_{s \in \mathcal{H}_{k^*}} \left(f(s) + \frac{1}{P} \right) && \text{(by definition of } f\text{)} \\
&\leq \sum_{s \in \mathcal{H}_{k^*}} \left[I_s \left(1 + \frac{1}{P} \right) + \frac{1}{P} \right] && \text{(by assumption)} \\
&\leq 1 + \sum_{s \in \mathcal{H}_{k^*}} \left[I_s \left(1 + \frac{1}{P} \right) \right] && (|\mathcal{H}_{k^*}| \leq P) \\
&\leq (I_g - 1) \left(1 + \frac{1}{P} \right) + 1 \\
&\leq I_g \left(1 + \frac{1}{P} \right),
\end{aligned}$$

where the third inequality is because k^* is not a descendant of any $s \in \mathcal{H}_{k^*}$, and thus

$$\sum_{s \in \mathcal{H}_{k^*}} I_s \leq I_g - 1.$$

Therefore, we show Eq. (50) by induction. □

E.5.2 The Counterpart of Eq. (11) for DOMWU

With Lemma 22, we can prove the following lemma, the counterpart of Eq. (11) for DOMWU.

Lemma 24. *Under the uniqueness assumption, there exists a constant $C_{14} > 0$ that depends on the game, η, and $\widehat{\boldsymbol{z}}_1$ such that for any $t \geq 1$, DOMWU with step size $\eta \leq \frac{1}{8P}$ guarantees*

$$D_{\Psi^{dil}_{\boldsymbol{\alpha}}}(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) + D_{\Psi^{dil}_{\boldsymbol{\alpha}}}(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t) \geq C_{14} D_{\Psi^{dil}_{\boldsymbol{\alpha}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1})$$

as long as $\max\{\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_t\|_1, \|\boldsymbol{z}^* - \boldsymbol{z}_t\|_1\} \leq \frac{\eta^2 \xi \epsilon_{dil}^2}{40P^3\|\boldsymbol{\alpha}\|_\infty}$.

Proof. We define $\alpha_{\min} = \min_{h \in \mathcal{H}^{\mathcal{Z}}} \alpha_h > 0$. Note that

$$\begin{aligned}
&D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) + D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}_t, \widehat{\boldsymbol{z}}_t) \\
&= \sum_{g \in \mathcal{H}^{\mathcal{Z}}} \alpha_g \cdot \widehat{z}_{t+1,\sigma(g)} \text{KL}(\widehat{\boldsymbol{q}}_{t+1,g}, \boldsymbol{q}_{t,g}) + \alpha_g \cdot z_{t,\sigma(g)} \text{KL}(\boldsymbol{q}_{t,g}, \widehat{\boldsymbol{q}}_{t,g}) && \text{(Eq. (24))} \\
&\geq \alpha_{\min} \sum_{g \in \mathcal{H}^{\mathcal{Z}}} \widehat{z}_{t+1,\sigma(g)} \text{KL}(\widehat{\boldsymbol{q}}_{t+1,g}, \boldsymbol{q}_{t,g}) + z_{t,\sigma(g)} \text{KL}(\boldsymbol{q}_{t,g}, \widehat{\boldsymbol{q}}_{t,g}) \\
&\geq \alpha_{\min} \epsilon_{\text{dil}} \sum_{g \in \mathcal{H},\ \sigma(g) \in \text{supp}(\boldsymbol{z}^*)} \text{KL}(\widehat{\boldsymbol{q}}_{t+1,g}, \boldsymbol{q}_{t,g}) + \text{KL}(\boldsymbol{q}_{t,g}, \widehat{\boldsymbol{q}}_{t,g}) && \text{(Lemma 20)} \\
&\geq \frac{\alpha_{\min}\epsilon_{\text{dil}}}{3} \sum_{i \notin \text{supp}(\boldsymbol{z}^*), p_i \in \text{supp}(\boldsymbol{z}^*)} \left(\frac{(\widehat{q}_{t+1,i} - q_{t,i})^2}{\widehat{q}_{t+1,i}} + \frac{(q_{t,i} - \widehat{q}_{t,i})^2}{q_{t,i}} \right) \\
& && \text{([Wei et al., 2021, Lemma 18])} \\
&\geq \frac{\alpha_{\min}\epsilon_{\text{dil}}}{4} \sum_{i \notin \text{supp}(\boldsymbol{z}^*), p_i \in \text{supp}(\boldsymbol{z}^*)} \left(\frac{(\widehat{q}_{t+1,i} - q_{t,i})^2}{\widehat{q}_{t,i}} + \frac{(q_{t,i} - \widehat{q}_{t,i})^2}{\widehat{q}_{t,i}} \right) && \text{(Lemma 18)} \\
&\geq \frac{\alpha_{\min}\epsilon_{\text{dil}}}{8} \sum_{i \notin \text{supp}(\boldsymbol{z}^*), p_i \in \text{supp}(\boldsymbol{z}^*)} \frac{(\widehat{q}_{t+1,i} - \widehat{q}_{t,i})^2}{\widehat{q}_{t,i}}. && (51)
\end{aligned}$$

Below we continue to bound $\sum_{i\notin \text{supp}(\boldsymbol{z}^*),p_i\in\text{supp}(\boldsymbol{z}^*)} \frac{(\widehat{q}_{t+1,i}-\widehat{q}_{t,i})^2}{\widehat{q}_{t,i}}$.

By the assumption, we have $\|\widehat{\boldsymbol{q}}_t - \boldsymbol{q}^*\|_1 \le P\|\widehat{\boldsymbol{z}}_t - \boldsymbol{z}^*\|_1/\epsilon_{\text{dil}}^2 \le \frac{\eta\xi}{10\alpha_i}$ and for index i such that $i \notin \text{supp}(\boldsymbol{x}^*)$ and $p_i \in \text{supp}(\boldsymbol{x}^*)$,

$$\sum_{j\in\Omega_{h(i)},j\notin\text{supp}(\boldsymbol{x}^*)} \widehat{q}_{t,j} \le \frac{\eta\xi}{10\alpha_i}. \tag{52}$$

Moreover, by Lemma 22, we have (denote $h = h(i)$)

$$\begin{aligned}
\widehat{q}_{t+1,i} &= \frac{\widehat{q}_{t,i}\exp(-\eta\widehat{L}_i/\alpha_i)}{\sum_{j\in\Omega_h}\widehat{q}_{t,j}\exp(-\eta\widehat{L}_j/\alpha_i)} \le \frac{\widehat{q}_{t,i}\exp(-\eta\widehat{L}_i/\alpha_i)}{\sum_{j\in\Omega_h\cap\text{supp}(\boldsymbol{z}^*)}\widehat{q}_{t,j}\exp(-\eta\widehat{L}_j/\alpha_i)} \\
&\le \frac{\widehat{q}_{t,i}\exp(-\eta(V_h^*+\frac{9\xi}{10})/\alpha_i)}{\sum_{j\in\Omega_h\cap\text{supp}(\boldsymbol{z}^*)}\widehat{q}_{t,j}\exp(-\eta(V_h^*+\frac{\xi}{10})/\alpha_i)} && \text{(Lemma 22)} \\
&= \frac{\widehat{q}_{t,i}\exp\left(-\frac{8}{10}\eta\xi/\alpha_i\right)}{\left(1-\sum_{j\in\Omega_h,j\notin\text{supp}(\boldsymbol{z}^*)}\widehat{q}_{t,j}\right)} \\
&\le \frac{\widehat{q}_{t,i}\exp\left(-\frac{8}{10}\eta\xi/\alpha_i\right)}{\left(1-\frac{\eta\xi/\alpha_i}{10}\right)} && \text{(Eq. (52))} \\
&\le \widehat{q}_{t,i}\left(1-\frac{1}{2}\frac{\eta\xi}{\alpha_i}\right). && \left(\tfrac{\exp(-0.8u)}{1-0.1u} \le 1-0.5u \text{ for } u\in[0,1]\right)
\end{aligned}$$

Rearranging gives

$$\frac{|\widehat{q}_{t+1,i}-\widehat{q}_{t,i}|^2}{\widehat{q}_{t,i}} \ge \frac{\eta^2\xi^2}{4\|\boldsymbol{\alpha}\|_\infty^2}\widehat{q}_{t,i} \ge \frac{\eta^2\xi^2}{8\|\boldsymbol{\alpha}\|_\infty^2}\widehat{q}_{t+1,i},$$

where the last step uses Lemma 18. The case for $\widehat{\boldsymbol{y}}_t$ is similar. Combining this with Eq. (51), we get

$$\begin{aligned}
D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\widehat{\boldsymbol{z}}_{t+1},\boldsymbol{z}_t) + D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}_t,\widehat{\boldsymbol{z}}_t) &\ge C_{12}' \sum_{i\notin\text{supp}(\boldsymbol{z}^*),p_i\in\text{supp}(\boldsymbol{z}^*)} \widehat{q}_{t+1,i} \\
&\ge C_{12}' \sum_{i\notin\text{supp}(\boldsymbol{z}^*)} z_{p_i}^*\widehat{q}_{t+1,i}
\end{aligned} \tag{53}$$

for some $C_{12}' > 0$. Now we combine two lower bounds of $D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\widehat{\boldsymbol{z}}_{t+1},\boldsymbol{z}_t) + D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}_t,\widehat{\boldsymbol{z}}_t)$. Using Lemma 12 and Eq. (53), we get

$$\begin{aligned}
&D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\widehat{\boldsymbol{z}}_{t+1},\boldsymbol{z}_t) + D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}_t,\widehat{\boldsymbol{z}}_t) \\
&= \frac{1}{2}\left(D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\widehat{\boldsymbol{z}}_{t+1},\boldsymbol{z}_t) + D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}_t,\widehat{\boldsymbol{z}}_t)\right) + \frac{1}{2}\left(D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\widehat{\boldsymbol{z}}_{t+1},\boldsymbol{z}_t) + D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}_t,\widehat{\boldsymbol{z}}_t)\right) \\
&\ge \frac{C_{12}}{2}\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_{t+1}\|_1^2 + \frac{C_{12}'}{2}\sum_{i\notin\text{supp}(\boldsymbol{z}^*)} z_{p_i}^*\widehat{q}_{t+1,i}.
\end{aligned} \tag{54}$$

Also note that by Lemma 17, we have

$$\begin{aligned}
D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}^*,\widehat{\boldsymbol{z}}_{t+1}) &\le \|\boldsymbol{\alpha}\|_\infty\left(\sum_{i\in\text{supp}(\boldsymbol{z}^*)}\frac{4P}{z_i^*}\frac{(z_i^*-\widehat{z}_{t+1,i})^2}{\widehat{q}_{t+1,i}} + \sum_{i\notin\text{supp}(\boldsymbol{z}^*)} z_{p_i}^*\widehat{q}_{t+1,i}\right) \\
&\le \frac{4\|\boldsymbol{\alpha}\|_\infty P}{\epsilon_{\text{dil}}^2}\left(\sum_{i\in\text{supp}(\boldsymbol{z}^*)}(z_i^*-\widehat{z}_{t+1,i})^2 + \sum_{i\notin\text{supp}(\boldsymbol{z}^*)} z_{p_i}^*\widehat{q}_{t+1,i}\right). && \text{(Lemma 20)}
\end{aligned}$$

Combining this with Eq. (54), we conclude that

$$\begin{aligned}
D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\widehat{\boldsymbol{z}}_{t+1},\boldsymbol{z}_t) + D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}_t,\widehat{\boldsymbol{z}}_t) &\ge \frac{\min\{C_{12},C_{12}'\}}{2}\left(\|\boldsymbol{z}^*-\widehat{\boldsymbol{z}}_{t+1}\|_1^2 + \sum_{i\notin\text{supp}(\boldsymbol{z}^*)} z_{p_i}^*\widehat{q}_{t+1,i}\right) \\
&\ge \frac{\epsilon_{\text{dil}}^2}{8\|\boldsymbol{\alpha}\|_\infty P}\min\{C_{12},C_{12}'\}\, D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}^*,\widehat{\boldsymbol{z}}_{t+1}),
\end{aligned}$$

which finishes the proof.

□

E.5.3 Proof of Theorem 7

With Lemma 24, we are ready to prove Theorem 7.

Proof of Theorem 7. Set $T_0 = \frac{64C'_{13}}{c^2}$, where $c = \frac{\eta^2\xi\epsilon_{\text{dil}}^2}{40P^3\|\boldsymbol{\alpha}\|_\infty}$. For $t \geq T_0$, we have using Theorem 21,

$$\begin{aligned}
\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_t\|_1^2 &\leq 2D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_t) \leq \frac{2C'_{13}}{T_0} \leq c^2, \\
\|\boldsymbol{z}^* - \boldsymbol{z}_t\|_1^2 &\leq 2\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_{t+1}\|_1^2 + 2\|\widehat{\boldsymbol{z}}_{t+1} - \boldsymbol{z}_t\|_1^2 \\
&\leq 4D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1}) + 4D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) \\
&\leq 64\Theta_{t+1} \leq \frac{64C'_{13}}{T_0} \leq c^2.
\end{aligned}$$

Therefore, when $t \geq T_0$, the condition of the second part of Lemma 24 is satisfied, and we have

$$\begin{aligned}
\zeta_t &\geq \frac{1}{2}D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) + \frac{1}{2}\zeta_t \\
&\geq \frac{1}{2}D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\widehat{\boldsymbol{z}}_{t+1}, \boldsymbol{z}_t) + \frac{C_{14}}{2}D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_{t+1}) && \text{(by Lemma 24)} \\
&\geq C_{15}\Theta_{t+1}.
\end{aligned}$$

for some constant $C_{15} > 0$. Therefore, when $t \geq T_0$, $\Theta_{t+1} \leq \Theta_t - \frac{15}{16}C_{15}\Theta_{t+1}$, which further leads to

$$\Theta_t \leq \Theta_{T_0} \cdot \left(1 + \frac{15}{16}C_{15}\right)^{T_0 - t} \leq \Theta_1 \cdot \left(1 + \frac{15}{16}C_{15}\right)^{T_0 - t} = D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1) \cdot \left(1 + \frac{15}{16}C_{15}\right)^{T_0 - t},$$

where the second inequality uses Eq. (43). The inequality trivially holds for $t < T_0$ as well, so it holds for all t. We finish the proof by relating $D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}^*, \boldsymbol{z}_t)$ and Θ_{t+1}. Note that by Lemma 17,

$$\begin{aligned}
D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}^*, \boldsymbol{z}_t)^2 &\leq \frac{16P\|\boldsymbol{\alpha}\|_\infty^2}{\epsilon_{\text{dil}}^4}\|\boldsymbol{z}^* - \boldsymbol{z}_t\|_1^2 \\
&\leq \frac{32P\|\boldsymbol{\alpha}\|_\infty^2}{\epsilon_{\text{dil}}^4}\left(\|\boldsymbol{z}^* - \widehat{\boldsymbol{z}}_{t+1}\|_1^2 + \|\widehat{\boldsymbol{z}}_{t+1} - \boldsymbol{z}_t\|_1^2\right) \\
&\leq \frac{1024P\|\boldsymbol{\alpha}\|_\infty^2}{\epsilon_{\text{dil}}^4}\Theta_{t+1}.
\end{aligned}$$

Therefore, we conclude

$$D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}^*, \boldsymbol{z}_t) \leq \sqrt{\frac{1024P\|\boldsymbol{\alpha}\|_\infty^2}{\epsilon_{\text{dil}}^4}\Theta_{t+1}} \leq \sqrt{\frac{1024P\|\boldsymbol{\alpha}\|_\infty^2 D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1)}{\epsilon_{\text{dil}}^4}}\left(1 + \frac{15}{16}C_{15}\right)^{\frac{T_0 - t - 1}{2}},$$

which finishes the proof by setting

$$C_3 = \sqrt{\frac{1024P\|\boldsymbol{\alpha}\|_\infty^2 D_{\Psi_{\boldsymbol{\alpha}}^{\text{dil}}}(\boldsymbol{z}^*, \widehat{\boldsymbol{z}}_1)}{\epsilon_{\text{dil}}^4}}\left(1 + \frac{15}{16}C_{15}\right)^{\frac{T_0 - 1}{2}}, \ C_4 = \left(1 + \frac{15}{16}C_{15}\right)^{\frac{1}{2}} - 1.$$

□

E.6 Remarks on DOGDA

In this subsection, we discuss the technical difficulties to get a convergence rate for DOGDA. This is challenging even if we assume the uniqueness of the Nash equilibrium. From the analysis of VOMWU and DOMWU, we can see that Lemma 19 and Lemma 20 play an important role. The lemmas lower bound $\widehat{z}_{t,i}$ with some game-dependent constants for i in supp($\boldsymbol{z}^*$), and the proofs are based on the observation that $D_{\Psi^{\text{van}}}(\boldsymbol{z}^*, \boldsymbol{z})$ and $D_{\Psi^{\text{dil}}_{\boldsymbol{\alpha}}}(\boldsymbol{z}^*, \boldsymbol{z})$ approach infinity as z_i approaches zero for $i \in \text{supp}(\boldsymbol{z}^*)$. This property of the entropy regularizers, however, does not hold for the dilated Euclidean regularizer $\Phi^{\text{dil}}_{\boldsymbol{\alpha}}$ in general. Lower bounding $\widehat{z}_{t,i}$ for DOGDA could be possible when $\widehat{\boldsymbol{z}}_t$ is sufficiently close to $\boldsymbol{z}^*$. For example, when

$$\|\widehat{\boldsymbol{z}}_t - \boldsymbol{z}^*\| \leq \frac{1}{2} \min_{i \in \text{supp}(\boldsymbol{z}^*)} z_i^*,$$

we can lower bound $\widehat{z}_{t,i}$ by $\frac{1}{2}\min_{i\in\text{supp}(\boldsymbol{z}^*)} z_i^*$ for $i \in \text{supp}(\boldsymbol{z}^*)$. This must happen when t is large by Theorem 4, but the entire analysis will then depend on a potentially large "asymptotic" constant. Therefore, even though we know that asymptotically, DOGDA has linear convergence, getting a concrete rate as VOMWU and DOMWU is still an open question.

Another direction is to follow the analysis of VOGDA, which gives a linear convergence rate by Corollary 5. However, in the analysis of VOGDA, Wei et al. [2021] implicitly use the fact that Φ^{van} is β-smooth, that is,

$$D_{\Phi^{\text{van}}}(\boldsymbol{z}, \boldsymbol{z}') \leq \frac{\beta}{2}\|\boldsymbol{z} - \boldsymbol{z}'\|^2,$$

for some $\beta > 0$. In fact, Φ^{van} is 1-smooth. This property, unfortunately, does not hold for $\Phi^{\text{dil}}_{\boldsymbol{\alpha}}$. We believe one can still show that $\Phi^{\text{dil}}_{\boldsymbol{\alpha}}$ is β-smooth for some game-dependent β once $\widehat{\boldsymbol{z}}_t$ is sufficiently close to $\boldsymbol{z}^*$. However, this again involves the asymptotic result in Theorem 4 and may prevent us from getting a concrete rate. In summary, using the existing techniques, we met some difficulties to obtain a concrete convergence rate for DOGDA. However, DOGDA performs well in the experiments. Moreover, it reduces to VOGDA in the normal-form games and achieves linear convergence in this case. Therefore, we still believe that it is a promising direction to get a (linear) convergence rate for DOGDA in theory.