Second-Order Mirror Descent: Convergence in Games Beyond Averaging and Discounting

Bolin Gao and Lacra Pavel

Abstract—In this article, we propose a second-order extension of the continuous-time game-theoretic mirror descent (MD) dynamics, referred to as MD2, which provably converges to mere (but not necessarily strict) variationally stable states (VSS) without using common auxiliary techniques, such as time-averaging or discounting. We show that MD2 enjoys no-regret as well as an exponential rate of convergence toward strong VSS upon a slight modification. MD2 can also be used to derive many novel continuous-time primal-space dynamics. We then use stochastic approximation techniques to provide a convergence guarantee of discrete-time MD2 with noisy observations toward interior mere VSS. Selected simulations are provided to illustrate our results.

Index Terms—Autonomous system, game theory, gradient methods, multi-agent systems, nonlinear dynamical systems, reinforcement learning.

I. INTRODUCTION

A CENTRAL problem of multiagent online learning is the design of adaptive policies which a set of subscribing agents (or players) can utilize to arrive at a desired collective outcome. These adaptive policies can typically be viewed as the following iterative process. An information processing step, whereby game-relevant information is made available to the player and then collected for processing, followed by a decision step, whereby the processed information is converted into the next strategy. A unified mathematical codification of this two-step process that emerged in recent years is referred to as mirror descent (MD), first studied in [1] and subsequently in [2], [3], [4], [5], [6], [7], [8], and [9] in the context of optimization.

For games, mirror descent is often analyzed as a set of ordinary differential equations (ODEs), referred to as the continuous-time MD dynamics, which can be seen as the multiagent extension of [3, p. 87]. MD is also referred to as dual averaging (DA) in continuous games [10], [11] and follow-the-regularized-leader (FTRL) or exponential learning in finite games [12], [13], [14]. The core appeal of MD lies in its flexibility in adapting to a wide variety of problem settings as well as its rich theoretical properties. In particular, MD is known to converge to the Nash equilibrium (NE) under a trio of strict assumptions.

1) The game is strictly monotone, i.e., the pseudogradient of the game is a strictly monotone operator [15]. This class captures games that admit strictly concave potential functions, as well as saddle-point problems with strictly convex, strictly concave saddle functions, a special case of diagonally strict concavity of [16].

2) The NE is strictly variationally stable (VS) (in the sense of [10]). The class of games with a strictly VS NE (also known as strictly VS games [17]) contains the class of strictly monotone games, as well as strictly coherent saddle point problems [18].

3) The NE is strict [19], which coincides with a locally strictly VS NE in finite games [10].

Despite these theoretical guarantees, the applicability of MD and its variants remains limited, as many games found in practice do not conveniently exhibit these strict properties. In other words, MD does not converge in many games, most notably in zero-sum (ZS) finite games that feature a unique interior NE [14]. Hofbauer and Sandholm [20] showed that no two-player ZS finite game is strictly monotone. Furthermore, strictness is also intimately related to the uniqueness of the game equilibrium. For instance, a strictly monotone game admits a unique NE [15]. In practice, however, many games exhibit a convex set of equilibria as opposed to a unique one. This means strictness imposes serious constraints on the problem parameters. Moreover, many dynamics or algorithms that are directly derived from MD (through discretization or otherwise) share similar limitations in that some type of strictness needs to be assumed to ensure convergence, e.g., [10], [11], [12], [13], [17], [21], [22], and [23].

To achieve convergence beyond strictness in a continuous-time setup, the existing literature typically relies on two primary approaches: averaging and discounting. The first method utilizes the time-averaged (ergodic) strategies generated by MD as opposed to the actual strategies [12]. However, this approach has two drawbacks. First, the players need to combine their time-averaged strategies in order to recover the NE, which is unrealistic in noncooperative settings. Second, time-averaging could fail outside of ZS games [24]. Another method is via a discounting procedure [25], [26], [27], which is conceptually related to the weight decay method of [28]. Discounting can be seen as a regularization technique that makes use of a strictly convex function to offset poor game properties, such as the lack of strict monotonicity. However, in general, discounting cannot yield exact convergence [25], [26], [27]. Since averaging and discounting are two of the most common methods for improving the convergence properties of continuous-time MD, yet both

Manuscript received 18 November 2021; revised 21 November 2021, 25 June 2022, and 12 January 2023; accepted 29 June 2023. Date of publication 4 July 2023; date of current version 29 March 2024. This work was supported by a grant from NSERC and Huawei Technologies Canada. Recommended by Associate Editor S. Grammatico. (Corresponding author: Bolin Gao.)

The authors are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: bolin.gao@mail.utoronto.ca; pavel@control.toronto.edu).

Digital Object Identifier 10.1109/TAC.2023.3291953

© 2023 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/
have been met with challenges, therefore it is natural for us to set our sights on alternative approaches.

**Contributions:** In this work, in a game-theoretic setup, we propose a second-order variant of the continuous-time MD (the first-order of which was studied in [2], [3], [11], [12], [13], and [14]), which we refer to as MD2. Second-order means that the set of ODEs is second-order in time, thus MD2 can be seen as a dual-space formulation of the HB method [29] and is closely related to the class of nth order discounted game dynamics of [26], adapted here in a more general game setting. Like MD, we show that MD2 satisfies the basic assumption of no-regret. However, unlike MD, MD2 converges without using averaging even when the NE is not strictly VS. Moreover, unlike the discounted MD [27], which partially overcomes the convergence issue of MD at the cost of inexact convergence, MD2 converges exactly. Exact convergence via a primal-space, second-order pseudogradient-type dynamics has been achieved in (merely) monotone games [30]. We provide a dual-space generalization of this result, beyond the monotone game setting, into the case where the equilibria are merely VS. Furthermore, MD2 can exactly recover the unconstrained, primal-space dynamics of [30] in the full-information case. Finally, we use our continuous-time convergence results to guide the proof of convergence of a discrete-time MD2. In gist, we show that higher order dynamics can converge toward Nash solutions with more general stability properties.

**Related works:** This article incorporates several ideas from the variational stability, higher order dynamics, and stochastic (semibandit) Nash-seeking literature. The concept of a variationally stable state (VSS) has roots in evolutionary game theory, and it is motivated by and analogous to that of an evolutionarily stable state (ESS) [31]. Namely, an ESS is a locally strict VSS for games with one or more populations of agents, also known as population games (similarly, a global ESS or GESS is a globally strict VSS) [20]. Many evolutionary game dynamics, such as the replicator and projection dynamics, have been shown to converge toward NEs of strictly stable games, which are GESS [21], [20]. In contrast to ESS, VSS is defined for the more general class of continuous games [10]. The literature on strict VSS seeking without structural property, such as strict monotonicity, is relatively recent [10], [11], [13], [17], [22]. Mertikopoulos and Staudigl [11] showed that DA converges toward a globally strict VSS. The authors in [17] and [22] studied online gradient descent (discrete MD with projection map) where the feedback is received asynchronously between agents.

In contrast to a strict VSS, the class of mere VSS is motivated by the more general notion of a neurally stable state (NSS) and global NSS (GNSS) [32]. Comparatively speaking, the set of literature that deals with convergence toward a mere VSS is fewer, especially in the absence of structural assumptions, such as the game being merely monotone (also known as stable game [20]). In a population game setup, Hofbauer and Sandholm [20] showed that the best response, integrable excess payoff, and impartial pairwise comparison dynamics converge globally to the set of NEs in (merely) stable games, which are GNSS. With a few exceptions, such as [18], the current set of literature on the (nonergodic) convergence toward a mere VSS outside of a finite/population game setup typically also assumes monotonicity of the game [33], [30], [34]. In contrast to the abovementioned references, in this work we generalize the convergence behavior of MD to nonstrict NE (i.e., a mere VSS) without any structural assumption, through the use of higher order augmentation.

Higher order dynamics was pioneered by Polyak [29], whose algorithm later became widely known as the heavy-ball method (HB) [29]. HB was recognized as an analogue of a second-order ODE upon its inception and has since been heavily investigated in the optimization literature [35] and arguably forms the backbone of deep learning, where nonconvexity and local minima dominate [36]. Inspired by [29], the authors of [37] studied a second-order dynamics for continuous concave games, which can be seen as a second-order extension of the projection dynamics of [38] and showed that their dynamics converge whenever there exists a Lipschitz and bounded potential function. In a finite game setup, Laraki and Mertikopoulos [39] studied nth-order variant of exponential learning [12], whereby the payoff is processed via successive integration, and showed that these higher order learning schemes can achieve faster rate of elimination of iteratively dominated solutions and convergence toward strict NE. An important connection between higher order dynamics and the stability of an NE was established by Arslan and Shamma [40], which showed that the addition of an anticipatory process to first-order gradient-play and replicator dynamics (RD) can result in local convergence toward an interior (nonstrict) NE, despite the dynamics themselves being uncoupled. For continuous-time equilibrium seeking via a higher order augmentation in the dual-space, the closest work related to ours is the second-order discounted exponential learning scheme in [26], which overcomes the nonconvergence toward interior NE of [39] at the cost of inexact convergence. It was found empirically in [26] that such higher order augmentation improves the convergence property of its first-order counterpart and converges in nonstrictly monotone games where the first-order fails. We improve [26] by generalizing these results from finite games to continuous games and go beyond a monotone game setup while simultaneously achieving exact convergence. Exact convergence (also known as last iterate convergence) via a primal-space, second-order pseudogradient dynamics was achieved in merely monotone games [30], which we also generalize, by stripping away monotonicity assumptions and showing that MD2 encapsulates the dynamics of [30] as a special case.

In the discrete-time, semibandit setup, whereby each player receives a (possibly) noise-corrupted version of the pseudo-gradient, the closest work related to ours is [10] studied the first-order MD, known as DA therein, and showed that under imperfect gradient setup, DA converges to a globally VS NE under diminishing step-sizes and its ergodic average converges to the set of NE in two-player ZS games. Several variations to DA have been studied in the semibandit literature, whereby their convergence is toward either strict (strong) VSS or strict NE. For instance, in finite games, Cohen et al. [41] studied a Hedge-variant of exponential weight algorithm (HEDGE) and provided convergence when the NE is strongly VS with respect to the $L^1$ norm. In a similar setup, Couchedey et al. [25] studied a discounted variant of HEDGE for potential games. A similar analysis was also performed by [23] for FTRG in finite games. Kashli et al. [42] provided several primal-space algorithms for merely monotone games. Mertikopoulos et al. [18] studied the optimistic MD algorithm, which converges in coherent games, i.e., two-player games that possess mere VSS. Compared to optimistic MD, discrete-time MD2 only requires one mirror projection instead of two prox projections. Furthermore, optimistic MD cannot converge in the presence of noise [18] unless the game is strictly coherent (a stronger assumption), whereas the discrete-time MD2 can converge under noisy conditions.

**Organization:** We provide the background materials in Section II. Section III reviews several basic properties of the first-order MD. We discuss the convergence properties of MD2 in
Section IV and provide the rate of convergence as well as regret minimization guarantee in Section V. We derive associated primal-space dynamics in Section VI. Section VII discusses MD2 in discrete-time with noisy observations. Simulations are presented in Section VIII. Finally, Section IX concludes this article. For readability, all proofs are relegated to the Appendix. The main notations used in this article are provided in Table I.

II. BACKGROUND

The following material is drawn from [43] and [44]. Those for game theory are drawn from [15], [45], and [46].

A. Sets and Vectors

Given a convex set \( C \subseteq \mathbb{R}^n \), its (relative) interior is denoted as \( \text{rint}(C) \). The nonnegative orthant of \( \mathbb{R}^n \) is \( \mathbb{R}^n_+ \), and the positive orthant as \( \mathbb{R}^n_+ \). A column vector in \( \mathbb{R}^n \) is denoted as \( x = (x_1, \ldots, x_n) \). 1 and 0 denote the column vector all ones and all zeros. \( I_{n \times n} \) and \( O_{n \times n} \) denote the \( n \times n \) identity and zero matrices. Subscript is omitted when the dimensionality is unambiguous. Suppose \( n \) is a natural number, then \( [n] := \{1, \ldots, n\} \).

B. Convex Functions and Duality

Let \( M \) be endowed with norm \( \| \cdot \| \) and dot product \( \langle \cdot, \cdot \rangle \). An extended real-valued function is a function \( f \) that maps from \( M \) to \( (-\infty, \infty] \). The effective domain of \( f \) denoted \( \text{dom}(f) = \{ x \in M : f(x) < \infty \} \). A function \( f : M \rightarrow (-\infty, \infty] \) is proper if \( f(x) \neq -\infty \) \( \forall x \) and there exists at least one \( x \in M \) such that \( f(x) < \infty \). f is closed if its epigraph is closed. A function \( f : M \rightarrow (-\infty, \infty] \) is supercoercive if \( \lim_{\|x\| \to \infty} f(x) / \|x\| \to \infty \). The indicator function over \( C \) is denoted by \( \delta_C \). The normal cone of \( C \) is defined as \( N_C(x) = \{ v \in \mathbb{R}^n : \langle v, y-x \rangle \leq 0 \ \forall y \in C \} \). The tangent cone of a nonempty convex set \( C \) at \( x \in C \) is \( T_{x}C = \bigcup_{y \in C} \lambda \cdot (C-x) \) and equals \( \emptyset \) for all \( x \notin C \). Id denotes the identity function.

Recall that \( \pi_C(x) = \argmin_{y \in C} \| y-x \|_2^2 \) is the Euclidean projection of \( x \) onto \( C \), where we refer to \( \pi_C \) as the Euclidean projector. Let \( \partial f(x) \) denote the set of subgradients of \( f \) at \( x \) and \( \nabla f(x) \) the gradient of \( f \) at \( x \). For differentiation on the boundary of a closed set \( C \), in lieu of the subgradient, we can also assume \( f \) is defined and differentiable on an open set containing \( C \). Given \( f \), the function \( f^* : M^* \rightarrow [-\infty, \infty] \) defined by \( f^*(z) = \sup_{x \in M} [x^T z - f(x)] \), is called the conjugate function of \( f \), where \( M^* \) is the dual space of \( M \), endowed with the dual norm \( \| \cdot \|_{*} \). \( f^* \) is closed and convex if \( f \) is proper. By the conjugate subgradient theorem [43], suppose \( f : \mathbb{R}^n \rightarrow (-\infty, +\infty) \) is proper, closed and convex and \( f^* \) is its Fenchel conjugate, then for any \( x, z \in \mathbb{R}^n \), \( x^T z = f(x) + f^*(z) \iff z \in \partial f(x) \iff x \in \partial f^*(z) \). The Bregman divergence of a proper, closed, convex function \( f \) is \( D_f : \text{dom}(f) \times \text{dom}(\partial f) \rightarrow [0, \infty) \), \( D_f(x, y) = f(x) - f(y) - \langle y-x, f^*(y) \rangle \), \( f \in \text{dom}(f) \). Given a vector-valued function \( F \), the Jacobian of \( F \) is denoted as \( J_F \).

C. N-Player Concave Games

Let \( G = (N, \{\Omega^p\}_{p \in N}, \{U^p\}_{p \in N}) \) be a game, where \( N = \{1, \ldots, N\} \) is the set of players, \( \Omega^p \subseteq \mathbb{R}^n \) is the set of player \( p \)'s strategies (actions). We denote the strategy (action) set of player \( p \)'s opponents as \( \Omega - \{p\} \subseteq \mathbb{R}^n \), \( \Omega^p(x) = \{ x \} \) and \( \Omega^p(x) = \{ x \} \) as player \( p \)'s payoff function, where \( x = (x^p, x^{-p}) \in \Omega^p \) is the action profile of all players, and \( x^p \in \Omega^p \) is the action of player \( p \). We also denote \( x = (x^p, x^{-p}) \) where \( x^p \in \Omega^p \) is the action profile of all players except \( p \). For differentiability purposes, we assume that there exists some open set, on which \( U^p \) is defined and continuously differentiable, such that it contains \( \Omega \).

Assumption 1: For all \( p \in N \), \( \Omega^p \) is a nonempty, closed, convex set, \( U^p : \Omega^p \rightarrow \mathbb{R}^n \) is (jointly) continuous in \( x = (x^p, x^{-p}) \). \( U^p(x^p, x^{-p}) \) is concave and continuously differentiable (\( C^1 \)) in \( x^p \) for all \( x^p \in \Omega^p \).

Under Assumption 1, we refer to \( G \) as a (continuous) concave game. Given \( x^{-p} \in \Omega^p \), each agent \( p \in N \) aims to find the solution of the following optimization problem:

\[
\begin{align*}
\text{maximize} & \quad U^p(x^p; x^{-p}) \quad \text{subject to} \quad x^p \in \Omega^p. \\
\text{subject to} & \quad x^p \in \Omega^p. \\
\text{subject to} & \quad \forall p \in N. \\
\text{subject to} & \quad \forall x \in \Omega. \\
\text{subject to} & \quad \forall x \neq x^* \in \Omega. \\
\text{subject to} & \quad \forall x \in \Omega. \\
\end{align*}
\]

(1) A profile \( x^* = (x^p)^{p \in N} \) is an NE if

\[
U^p(x^p; x^{-p}) \geq U^p(x^p; x^*-p^*) \quad \forall x^p \in \Omega^p \quad \forall p \in N. 
\]

(2) i.e., no player can increase its payoff by unilateral deviation.

Remark 1: Under Assumption 1, existence of an NE is guaranteed for bounded \( \Omega^p \) [49, Th. 4.4]. When \( \Omega^p \) is closed but not bounded, the existence of an NE is guaranteed under the additional assumption that \( -U^p \) is coercive in \( x^p \), that is, \( \lim_{\|x^p\| \to \infty} -U^p(x^p; x^{-p}) = +\infty \) for all \( x^p \in \Omega^p \). By [50, Prop. 1.4.2], \( x^* \in \Omega \) is an NE iff

\[
(x - x^*)^T U(x^*) \leq 0 \quad \forall x \in \Omega. 
\]

(4) Equivalently, \( x^* \) is a solution of the Stampacchia variational inequality VI(\( \Omega, -U \)), [15]. We say that an NE is globally strict if the inequality of (4) is held strictly for all \( x^* \neq x \) [21].

D. Monotonicity

A general class of games in which many dynamics are guaranteed to converge is the class of monotone games, also known

| Symbol | Definition |
|--------|------------|
| \( x(t)/z(t) \) | Strategy/at time t/at time k |
| \( z(t)/\Omega \) | Dual aggregate/at time t/at time k |
| \( \xi(x)/\Xi \) | Primal aggregate/at time t/at time k |
| \( \langle U^p \rangle \) | Payoff/partial-gradient/pseudo-gradient |
| \( J_U \) | Jacobian of U/symmetric game Jacobian of U |
| \( \partial f/\partial f^* \) | Regularizer/\( \partial f^* \)/convex conjugate of \( f \) |
| \( C_p/C_\mu/J_C \) | Player p's/overall mirror map/Jacobian of \( C \) |
| \( \gamma, \xi, \alpha, \beta \) | Parameters associated with MD/MD2/DA2 |
| \( 1/\mathcal{O} \) | Identity matrix/zero matrix |

TABLE I

LIST OF MAIN NOTATIONS USED IN THIS ARTICLE
as stable games [20], [47] or dissipative games [48]. We contrast known definitions of monotone games in the literature.

**Definition 1:** The game $G$ is as follows.

1. $\eta$-strongly monotone if $(U(x) - U(x'))^\top (x - x') \leq -\eta \|x - x'\|^2_2 \forall x, x' \in \Omega$, for some $\eta > 0$.
2. Strictly monotone if $(U(x) - U(x'))^\top (x - x') < 0 \forall x \in \Omega \setminus \{x'\}$, with equality if and only if $x = x'$. 
3. (Merely) Monotone if $(U(x) - U(x'))^\top (x - x') \leq 0 \forall x, x' \in \Omega$.
4. $\mu$-weakly monotone if $(U(x) - U(x'))^\top (x - x') \leq \mu \|x - x'\|^2_2 \forall x, x' \in \Omega$, for some $\mu > 0$.

Strictly and strongly monotone games have been extensively investigated in the literature, at least dating from [16]. Merely monotone games have been studied in [20], [26], [27], [30], [33], [42], and [34]. Weakly monotone games were considered in [26], [49], and [50]. When $U$ is $C^1$, there exists a natural characterization in terms of definiteness of its Jacobian [50, p. 155, Prop. 2.3.2]. Note that strictly monotone games can have at most one NE, whereas NEs can form a nonsingleton convex set in merely monotone games [15].

**E. Variational Stability**

Although many classical examples of games satisfy monotonicity properties [20], these conditions may not hold in more complex scenarios. A recent line of research has started to relax monotonicity notions from a game to that of an equilibrium via the notion of a VS equilibrium or state [10].

**Definition 2:** An NE $x^* \in \Omega$ is as follows.

1. $\eta$-strongly VS if $U(x)^\top (x - x^*) \leq -\eta \|x - x^*\|^2_2 \forall x \in \Omega$, for some $\eta > 0$.
2. Strictly VS if $U(x)^\top (x - x^*) \leq 0 \forall x \in \Omega$, with equality if and only if $x = x^*$.
3. (Merely) VS if $U(x)^\top (x - x^*) \leq 0 \forall x \in \Omega$.
4. $\mu$-weakly VS if $U(x)^\top (x - x^*) \leq \mu \|x - x^*\|^2_2 \forall x \in \Omega$, for some $\mu > 0$.

If a condition (i)–(iv) holds on $D \subset \Omega$, then the definition is said to hold locally. Otherwise is said to hold globally.

**Remark 2:** We refer to $x^*$ as a globally strong/strict/merely/weak VS if it satisfies one of the corresponding VS notions in Definition 2 on all of $\Omega$; $x^*$ is a locally strong/strict/merely/weak VS otherwise. For practical reasons, we will informally refer $\mu$-weak VS with a small $\mu$ as a nearly merely VS, i.e., a weak VS within a small distance away from becoming a mere VS.

Globally mere VS are the solutions to Minty variational inequality [15]. In a population game context, the set of globally mere VS is called GNSS and (the unique) globally strict VS is called GESS [20], [21], [47]. Strict VS was extended to a local/global setwise definition in [10]. Zhou et al. [22] introduced a slight variation of strict VS called $\lambda$-VS. Mazumdar et al. [51] studied a local version of strict VS under the name locally asymptotically stable differential NE, which requires twice continuous-differentiability of the payoff functions. Strong VS (Definition 2(i)) was also studied in [10]. Cohen et al. [41] studied a variant of the strong VS with $L^1$ norm.

It can be seen that any NE of the strongly/strictly/merely/weak monotone game is a strong/strict/merely/weak VS. In particular, the nonempty set of NE coincides with the set of mere versus VS for merely monotone concave games [2, Th. 2.2]. When $\text{int}(\Omega) \neq \emptyset$ and $x^*$ is an interior NE, i.e., $U(x^*) = 0$, the condition for VS recovers the condition associated with monotone at $x^* = x^*$.

Hence, VS can be seen as a type of pointwise monotonicity and it is known to have a second-order characterization similar to that for monotone games, but specifically at the NE, see [10].

Recall that an NE $x^*$ is interior if it lies in the interior of $\Omega$, that is, $x^* \in \text{rint}(\Omega)$. Throughout this work, we make the following assumption.

**Assumption 2:** $G$ admits an interior mere VS.

**F. Second-Order Characterization of VS**

In practice, outside of monotone games, it is often difficult to verify that an NE is a particular type of VS directly through Definition 2. A more common approach to characterize the VS conditions of a solution is by looking at the symmetric game Jacobian whenever the pseudogradient $U$ is $C^1$, which is

$$
\mathcal{J}_U(x) := \frac{1}{2} (\mathbf{J}_U(x) + \mathbf{J}_U(x)^\top), x \in \Omega.
$$

The matrix $\mathcal{J}_U(x) \in \mathbb{R}^{n \times n}$ is symmetric and thus guaranteed to have real eigenvalues, which makes it amenable to analysis.

We now provide several sufficient conditions for verifying whether a VS $x^*$ is strict, mere or weak. This is performed by checking the definiteness of $\mathcal{J}_U(x)$ or $\mathcal{J}_U(x)^\top$. Our condition for strict VS is the same as that in [10]. Before proceeding, we say that the (symmetric) game Jacobian $\mathcal{J}_U(x) \in \mathbb{R}^{n \times n}$ is negative definite on $T_{\Omega}(x)$ if

$$
y^\top \mathcal{J}_U(x)y < 0 \forall x \in \Omega \forall y \in T_{\Omega}(x), y \neq 0
$$

and negative semidefinite on $T_{\Omega}(x)$ if the preceding inequality is nonstrict. We use the shorthand notations $\mathcal{J}_U(x)^\top < 0$ and $\mathcal{J}_U(x) \preceq 0$, respectively. For $\mu > 0$, $\mathcal{J}_U(x) - \mu I$ is written as $\mathcal{J}_U(x) \preceq I$. Similar conventions for when $x = x^*$.

**Proposition 1:** Let $x^* \in \Omega$ be an NE of $G$. Suppose $U$ is continuously differentiable, and:

i) $\mathcal{J}_U(x) \prec_0 T_{\Omega}(x) \forall x \in \Omega$, then $x^*$ is globally strictly VS and isolated;
ii) $\mathcal{J}_U(x) \preceq_0 T_{\Omega}(x) \forall x \in \Omega$, then $x^*$ is globally merely VS;
iii) $\mathcal{J}_U(x) \preceq I$ $T_{\Omega}(x) \forall x \in \Omega$, then $x^*$ is globally $\mu$-weakly VS.

Suppose instead:

i') $\mathcal{J}_U(x^*) < 0$ on $T_{\Omega}(x^*)$, then $x^*$ is locally strictly VS and isolated;
ii') $\mathcal{J}_U(x^*) \preceq 0$ on $T_{\Omega}(x^*)$, then $x^*$ is locally merely VS;
iii') $\mathcal{J}_U(x^*) \preceq I$ on $T_{\Omega}(x^*)$, then $x^*$ is locally $\mu$-weakly VS.

**Remark 3:** These conditions can be verified by calculating the maximum eigenvalue of $\mathcal{J}_U(x)$ (respectively, $\mathcal{J}_U(x^*)$), where $\lambda_{\text{max}}(\mathcal{J}_U(x))$ (respectively, $\lambda_{\text{max}}(\mathcal{J}_U(x^*))$) denotes the (real) eigenvalue with the largest magnitude associated with an eigenvector in $T_{\Omega}(x)$ (respectively, $T_{\Omega}(x^*)$). When $\mathcal{J}_U(x)$ (respectively, $\mathcal{J}_U(x^*)$) is symmetric, then analysis can be performed directly on $\mathcal{J}_U$ without resorting to calculating $\mathcal{J}_U$.

**Remark 4:** Note that, even if $x^*$ is shown to be $\mu$-weak through Proposition 1(iii) or (iii’), it does not preclude the possibility that $x^*$ could in fact be strict or mere due to the looseness of the bound in the proof of Proposition 1. Furthermore, unlike strict VS, mere VS need not be isolated. The next example illustrates various notions of variational stability.

**Example 1:** (Monotone game with a mere VS) Every NE of a merely monotone game is globally mere VS. Perhaps
the simplest example of a merely monotone game is the so-called bilinear saddle point problem, whereby, suppose we have a saddle function, $f(x^1, x^2) = x^1 x^2, x^p \in \mathbb{R}$ for which we want to minimize in $x^1$ and maximize in $x^2$. We can cast this saddle point problem as a game by utilizing two payoff functions $U^1(x^1; x^2) = -x^1 x^2$ and $U^2(x^2; x^1) = x^1 x^2$. This game has a pseudogradient of $U(x) = (-x^1, x^2)$ and a game Jacobian $\mathbf{J}_U(x) = 0$. By Proposition 1(ii), the unique interior NE $x^* = (0, 0)$ is the globally mere VSS.

Example 2: (Nonmonotone potential game with a mere VSS) This example shows that the mere monotonicity enjoyed by Example 1 can be destroyed by the addition of another player, even when all the player's payoff functions remain linear in its own argument. Considered a three-player game, whereby

$$U^p(x^P; x^{-p}) = -x^p \sum_{q \neq p} x^q, \quad p \in \{1, 2, 3\}$$

and each $x^p \in [-1, 1]$, i.e., $\Omega = [-1, 1]^3$. For this game, the pseudogradient is $U(x) = (x^2 x^3, x^2 x^1, x^1 x^2)$, which shows that there exists an interior NE at $x^* = (0, 0, 0)$. The Jacobian, which is symmetric, is

$$\mathbf{J}_U(x) = -\begin{bmatrix} 0 & x^3 & x^2 \\ x^3 & 0 & x^1 \\ x^2 & x^1 & 0 \end{bmatrix}$$

hence, this game is nonmonotone in general and in fact possesses a nonconvex and nonconcave potential function $P(x) = -x^1 x^2 x^3$, i.e., $\nabla P = U$. Since $\mathbf{J}_U(x^*)$ is Osis negative semidefinite, hence by Proposition 1(ii), $x^*$ is locally merely VS. If we restrict the strategy set to $\Omega = [0, 1]^3$ instead, every NE $x^* \in \{x \in \Omega | x^p = 0, p, q \in \mathbb{N}, p \neq q\}$ is a globally mere VSS, as $U(x)^T (x - x^*) = -3 x^1 x^2 x^3 \leq 0 \forall x \in [0, 1]^3$.

Example 3: (Rock–paper–scissors (RPS) with nonnegative payoff for ties) Consider a two-player RPS game with $A$ and $B$ being the payoff matrices for player 1 and 2, respectively

$$A = \begin{bmatrix} \varsigma & -\ell & w \\ w & \varsigma & -\ell \\ -\ell & w & \varsigma \end{bmatrix}, \quad B = A^T$$

where, $\ell, w \geq 0$ are the values associated with a loss or a win and $\varsigma \in \mathbb{R}$ is the payoff of a tie. The strategy set associated with this game is the simplex $\Omega^p = \{x^p \in \mathbb{R}^n | \sum_{i=1}^n x_i^p = 1, x_i^p \geq 0\}$. The pseudogradient (as known as payoff vector $[20], [47]$) and the game Jacobian are as follows:

$$U(x) = \begin{bmatrix} O & A & x^1 \\ A & O & x^2 \\ x^2 & A & O \end{bmatrix}, \quad \mathbf{J}_U(x) = 1/2 \begin{bmatrix} O & A & A^T \\ A & O & A^T \\ A^T & A & O \end{bmatrix}.$$  

To simplify our analysis, consider an example where $\ell = 0$. In this game, $x^* = (x^p)^* \in \mathbb{N}, x^* = (1/3, 1/3, 1/3)$ is the unique interior NE for any $\varsigma \in [0, w], w > 0$. The eigenvalues of $\mathbf{J}_U(x)$ on $\mathbb{R}^3$ are $\{-\varsigma + w, \pm(\varsigma - w/2), \pm(\varsigma - w/2)\}$, hence this game is $\mu$-weakly merely monotone with $\mu = w/2$ whenever $\varsigma \in (w/2, w)$. When $\varsigma = w/2$, the game is merely monotone (Definition 1). Taken together, this means for $\varsigma \in (w/2, w), x^*$ is a unique, locally $\mu$-weak VS where $\mu = w/2$, whereas for $\varsigma = w/2, x^*$ is a unique, globally mere VS. Finally, note that when $\ell = w = \varsigma = 0$ (and hence both $A, B$ are zero matrices), every strategy is globally merely VS and forms a convex set (the strategy set itself).

### III. Review of First-Order MD Dynamics

We now describe the dynamic process that a group of agents utilizes in order to arrive at one of the equilibrium notions discussed in previous sections. One such general model is the MD dynamics. Intuitively, the family of MD dynamics ascribes an abstract behavior model to each player, which states that, upon receiving the partial-gradient of its payoff function, each player processes the partial-gradient information (typically via an aggregation), then converts the processed information into the next strategy. This process can be described by the following set of ODEs $[11], [12], [13], [14]$

$$\dot{z} = \gamma \nabla x U^p(x, z^{-p}) = \gamma U^p(x), \quad x^p = C_{\ell}^p(x^p) \quad (MD)$$

where, $x^p$ is referred to as score or dual variable, $\gamma > 0$ is a rate parameter, and $C_{\ell}^p : \mathbb{R}^n \to \mathbb{R}^n$ is referred to as the mirror map

$$C_{\ell}^p(x^p) = \arg \max_{y^p \in \mathbb{R}^n} \left[ y^p^T \pi^p - \epsilon d^p(y^p) \right], \epsilon > 0$$

where $\partial^p : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is assumed to be a closed, proper, and (at least) strictly convex function, referred to as a regularizer, where $d^p(\partial^p)$ is assumed to be a nonempty, closed, and convex set, which agrees with the strategy set $\mathbb{R}^n, \epsilon > 0$ is referred to as the regularization constant. The regularizer often satisfies one of the following distinct assumptions that ensure the existence of a unique solution of MD and (10).

Assumption 3: $\partial^p : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is closed, proper, convex, with $d^p(\partial^p) = 0$, nonempty, closed, and convex. In addition, $\partial^p$ is i) Legendre (i.e., strictly convex, steep, int$(\partial^p(0)) = \emptyset$) and supercoercive or ii) $\rho$-convex, $\rho > 0$.

We note that for $\partial^p$ to be steep, it means that $\|\nabla \partial^p(x^p)\| \to +\infty$, $\|x\|$, is the dual norm, whenever $\{x^p\}_k \in \mathbb{R}^n$, is a sequence in $\text{rint}(\partial^p(0)) = \text{rint}(\partial^p)$ converging to a point in the (relative) boundary. Let $\psi_p$ be the convex conjugate of $\psi_p$.(ii) Then by Lemma 3 and Lemma 2 (in the Appendix), under Assumption 3, $C_{\ell}^p = \nabla^p \psi_p$. When $\partial^p$ is steep, $C_{\ell}^p$ maps from $\mathbb{R}^n$ to all values in $\mathbb{R}^n$ except those at the boundary; $C_{\ell}^p$ could map to boundary points otherwise. We refer to $C_{\ell}^p$ as the mirror map induced by $\psi_p$ and Legendre or strongly induced (by $\psi_p$) when $\partial^p$ satisfies Assumption 2(i) or (ii). Note that the notions of Legendre and strongly convex need not be dichotomous. It is possible for a $C_{\ell}^p$ to be both Legendre and strongly induced. The following examples capture the three types of mirror maps that are extensively used in the literature $[2], [3], [11], [12], [13], [14], [17], [25], [26], [27]$.

Example 4: Let $\Omega^p = \mathbb{R}^n, \partial^p(x^p) = \frac{1}{2} \|x^p\|^2_2, \text{hence, it satisfies Assumption 3(i), (ii)}. In this case, $C_{\ell}^p = \mathbb{I}^d$ and it is both Legendre and strongly induced.

Example 5: Let $\Omega^p = \mathbb{R}^n, \partial^p(x^p) = \frac{1}{2} \|x^p\|^2_2$, where we assume $0 \log(0) = 0$. $\partial^p$ is 1-strongly convex in $\|\cdot\|_2$, hence it satisfies Assumption 3(ii). $C_{\ell}^p(x^p) = (\exp(\epsilon^{-1} x_i^p)(\sum_{i=1}^n \exp(\epsilon^{-1} x_i^p)))^{-1} \mathbb{1}_{\{n_p\}}$ is the softmax function.

Example 6: Let $\Omega^p \subset \mathbb{R}^n$ be a nonempty, compact, convex set, $\partial^p(x^p) = \frac{1}{2} \|x^p\|^2_2$, hence satisfies Assumption 3(ii) and is nonsteep. $C_{\ell}^p = \pi^p$ the Euclidean projection onto $\Omega^p$.

We can represent MD in a more compact stacked notation

$$\dot{z} = \gamma U(x), \quad x = C_{\ell}(z)$$

Observe that at rest, MD (11) satisfies

$$\mathbf{0} = U(\pi), \quad \pi = C_{\ell}^p(\pi), \quad \pi \in C_{\ell}^{-1}(\pi)$$
where, \( \mathbf{\tau} = (\mathbf{\tau}^p)_{p \in \mathcal{N}}, \mathbf{\tau} = (\mathbf{\tau}^p)_{p \in \mathcal{N}} \) are the rest points of MD and \( C_{\mathcal{N}}^{-1} = (C_p^p)^{-1})_{p \in \mathcal{N}} = (\partial \psi^p)_{p \in \mathcal{N}} \) is the inverse (or the preimage) of \( C_{\mathcal{N}} \). The rest condition (12) implies that \( \mathbf{\tau} \) is an interior NE. Hence, if a trajectory \( z(t) \) of MD comes to a rest, \( x(t) = C_{\mathcal{N}}(z(t)) \) converges to an interior NE. We note that the uniqueness of \( \mathbf{\tau} \) does not imply the uniqueness of \( \mathbf{\tau} \) unless \( C_{\mathcal{N}} \) is Legendre-induced (for which \( C_{\mathcal{N}} \) is one-to-one; this follows from Legendre theorem [52]). Hence, in general, the convergence of MD is to a set and not to an equilibrium and \( z(t) \) may continue to evolve even after \( x(t) \) has reached an equilibrium. Note that rest points are not the only game relevant solutions for which MD may converge to. As pointed out in [12], there are nonrest points that occur on the boundary of \( \Omega \) that are asymptotically stable.

Lemma 1: Let \( \mathcal{G} \) be a concave game with a globally strict VSS \( x^* \). Suppose that all players choose strategies according to MD (11). Let \( x(t) = (x^p(t))_{p \in \mathcal{N}} = C_{\mathcal{N}}(z(t)) \) be generated by (11) and \( C_{\mathcal{N}} = (C_p^p)_{p \in \mathcal{N}} \) be induced by \( \partial \psi^p \). For any \( \gamma, \epsilon > 0 \) and any \( x(0) = C_{\mathcal{N}}(z(0)) \in \Omega, z(0) \in \mathbb{R}^n \):

i) suppose \( \partial \psi^p \) satisfies Assumption 3(i) \( \forall p \), then \( x(t) = C_{\mathcal{N}}(z(t)) \) converges to \( x^* \), whenever \( x \) is interior;

ii) suppose \( \Omega^p \) is compact and \( \partial \psi^p \) satisfies Assumption 3(ii) \( \forall p \), then \( x(t) = C_{\mathcal{N}}(z(t)) \) converges to \( x^* \).

In general, MD does not converge beyond strict VSS, i.e., a mere VSS. This poses considerable limitations in practice, for instance, mere VSS are commonly found in ZS games. The authors in [20], [47] showed that every ZS finite game is merely (but not strictly) monotone, hence all of its NEs are mere (non-strict) VSS. A standard method to overcome nonconvergence to the NE in ZS games is to calculate a time-averaged (ergodic) trajectory \( x_{avg}(t) = t^{-1} \int_0^t x(\tau) d\tau \) in tandem with MD, that is

\[
\dot{z} = \gamma U(x), \quad x = C_{\mathcal{N}}(z), \quad x_{avg}(t) = t^{-1} \int_0^t x(\tau) d\tau \quad (MDA)
\]

for which the time-averaged strategy \( x_{avg} \) has been shown in many contexts to converge, e.g., [12]. The main critique of using MDA is that the actual strategies do not arrive at the NE in the long run, thereby making it unsuitable for online equilibrium seeking in the absence of a central planner or coordination between players. Furthermore, averaging may fail to converge outside of ZS games [24], which makes this approach vulnerable to parameter perturbation.

Another method for overcoming nonconvergence is through discounting, which was studied in [27]

\[
\dot{z} = \gamma(-z + U(x)), \quad x = C_{\mathcal{N}}(z) \quad (MDM)
\]

where, DMD stands for discounted MD. Compared to MD, an extra \(-z\) term is inserted in the \( z \) system, which translates into an exponential weighted decay (or discounting) term in the closed-form solution \( z(t) \). DMD has a connection with the so-called weight decay method in the machine learning literature [28], as \( z \in C_{\mathcal{N}}^{-1}(x) \) can be shown to be equivalent to a (usually non-Euclidean) regularization term, which interacts with the monotonicity of \( U \). While it is known that DMD could converge exactly in subclasses of finite games that exhibit symmetric interior NEs, (see [26]), in general, it cannot converge exactly to an NE, which also means that it cannot converge exactly to a VSS. This directs our attention to other methods, such as higher order augmentation [26], [37], [39], [40].

IV. SECOND-ORDER MD DYNAMICS

We now propose the second-order MD, which in terms of each player \( p \) appears as

\[
\begin{align*}
\dot{z}^p &= \gamma(U^p(x) - \alpha(x^p - \xi^p)), \\
\dot{\xi}^p &= \beta(x^p - \xi^p) \\
x^p &= C_p^p(z^p)
\end{align*}
\]

and in stacked notation

\[
\dot{z} = \gamma(U(x) - \alpha(x - \xi)), \quad \dot{\xi} = \beta(x - \xi), \quad x = C_{\mathcal{N}}(z) \tag{13}
\]

where, \( x = (x^p)_{p \in \mathcal{N}}, z = (z^p)_{p \in \mathcal{N}}, \xi = (\xi^p)_{p \in \mathcal{N}}, C_{\mathcal{N}} = (C_p^p)_{p \in \mathcal{N}}, U = (U^p)_{p \in \mathcal{N}} \). The rest point condition for MD2 is

\[
\dot{0} = U(\mathbf{\tau}), \quad 0 = \mathbf{\tau} - \mathbf{\xi}, \quad \mathbf{\tau} = C_{\mathcal{N}}(\mathbf{\tau}), \quad \mathbf{\tau} \in C_{\mathcal{N}}^{-1}(\mathbf{\tau}) \tag{14}
\]

which coincides with that of MD, i.e., \( \mathbf{\tau} = x^* \) is an interior NE.

We proceed to demonstrate that the primal aggregate \( \xi \) contributes to the convergence of MD2 beyond strict VSS and such convergence depends on both the property of the regularizer as well as the topological properties of the underlying strategy set.

Theorem 1: Let \( \mathcal{G} \) be a concave game and assume that every interior NE is globally merely VS. Suppose that all players choose strategies according to MD (13). Let \( x = (x^p(t))_{p \in \mathcal{N}} = C_{\mathcal{N}}(z(t)) \) be generated by (13) and \( C_{\mathcal{N}} = (C_p^p)_{p \in \mathcal{N}} \) be induced by \( \partial \psi^p \). Let \( x^* \) denote an interior mere VSS (possibly from a set of them). For any \( \alpha, \beta, \gamma, \epsilon > 0 \), \( x(0) = C_{\mathcal{N}}(z(0)) \in \Omega, z(0) \in \mathbb{R}^n \):

i) suppose \( \partial \psi^p \) is twice-continuously differentiable \((C^2)\) and satisfies Assumption 3(ii) \( \forall p \), then \( x(t) = C_{\mathcal{N}}(z(t)) \) converges to \( x^* \);

ii) assume \( \Omega^p \) is compact and \( \partial \psi^p \) satisfies Assumption 3(ii) \( \forall p \), then \( x(t) = C_{\mathcal{N}}(z(t)) \) converges to \( x^* \).

As a direct corollary of Theorem 1, MD2 converges to an interior mere NE of pseudomonotone games, and interior strict NE of quasimonotone games, see Corollary 1 in [64].

V. ADDITIONAL CONVERGENCE PROPERTIES OF MD2

In this section, we investigate two additional properties of MD2, namely, that of rate of convergence and regret minimization.

To account for the geometry of the problem, we provide a relative extension to the strong VSS.

Definition 3: Let \( h : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\} \) be any differentiable, convex function with domain \( dom(h) = \Omega \). Then, \( x^* \in \Omega \) is \( \eta \)-relatively strongly VS (with respect to \( h \)) if for all \( x \in \Omega \), \( U(x)^{-1}(x - x^*) \leq -\eta(D_h(x, x^*) + D_h(x^*, x)) \), for some \( \eta > 0 \), where \( D_h \) is the Bregman divergence of \( h \).

We note that Definition 3 is analogous to that of a relatively strongly monotone game, which was previously introduced in [53]. Following [29, Th. 4.4], it can be shown that MD converges to a relatively strongly VSS in \( \mathcal{O}(e^{-\gamma T \eta^{-1}}) \) given a strongly induced \( C_{\mathcal{N}} \) that is adapted to the geometry of this VSS. We now wish to provide a similar result for the convergence of MD2 toward a relatively strong VSS. However, exponential convergence does not follow from Theorem 1. Instead, we propose an augmented version of MD2 that exhibits exponential convergence

\[
\begin{align*}
\dot{z} &= (U(x) - \gamma \alpha(x - \xi)), \quad \dot{\xi} = \beta(x - \xi), \\
\dot{x} &= C_{\mathcal{N}}(z), \quad \dot{y} = -\eta e^{-\gamma t}
\end{align*}
\]

where, \( \gamma(0) > 0 \). Observe that MD2,\( \gamma \) is equivalent to the nonautonomous system

\[
\begin{align*}
\dot{z} &= (U(x) - e^{-\eta \gamma^{-1} t} \gamma(0) \alpha(x - \xi)), \quad \dot{\xi} = \beta(x - \xi), \\
\dot{x} &= C_{\mathcal{N}}(z)
\end{align*}
\]

(15)
whose rest point condition coincides with that of MD2.

Remark 5: From (15), MD2γ can be seen as MD with a vanishing perturbation \( g(t, z, \xi) = e^{-\nu t} \alpha(0)(x - \xi) \), which allows for the following simple learning interpretation. When the players are aware that the game being played has a \( \eta \)-strong VSS, they no longer bother with exploring the strategy space during the initial stages and instead discard the extra information represented by \( x - \xi \) exponentially fast.

Theorem 2: Let \( \mathcal{G} \) be a concave game with a unique interior \( \eta \)-strongly VSS relative to \( h(x) = \sum_{p \in N} \vartheta^p(x^p) \), denoted by \( x^* \), where \( x = (x^p(t))_{p \in N} = C_\xi(z(t)) \) is generated by MD2γ and \( C_\xi = (C_\xi^p)_{p \in N} = (\nabla \psi^p)_{p \in N} \).

Under this assumption, taking the time-derivative of \( x = C_\xi(z) \) shows that MD2 can be written as a pair of differential inclusions

\[
\dot{\xi} \in \beta(x - \xi), \quad \dot{x} = \gamma J_{C_\xi}(x)(U(x) - \alpha(x - \xi))
\]

where, \( \gamma > 0, \alpha, \beta \geq 0 \) and the Jacobian of \( C_\xi \) is

\[
J_{C_\xi}(z) := \text{blkdiag}(J_{C_\xi}^P(z^P))
\]

and each \( J_{C_\xi}(z) \) satisfies Assumption 3(ii). We note that the inclusions in (19) come from \( z \in C_\xi^{-1}(x) \). We can further express (19) in each \( p \) as

\[
\dot{\xi}^p \in \beta(x^p - \xi^p), \quad \dot{x}^p \in \gamma \nabla^2 \psi^p(x^p)(U^p(x) - \alpha(x^p - \xi^p)).
\]

In the following, we offer two general ways of rewriting (20) as a set of ODEs in the primal-space.

General Case: Our next result shows that finding the primal-space dynamics associated with MD2 generally amounts to evaluating the operator \( \nabla^2 \psi^p \circ \nabla \psi^p := \nabla^2 \psi^p(\nabla \psi^p) \).
which we refer to as the second-order natural gradient descent (NG2). NG2 is so named because in the optimization setup \((p = 1)\), for \(U = -\nabla f\), where \(f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}\) is some objective function and \(\alpha, \beta, \gamma > 0\), is analogous to the so-called natural gradient descent [54], \(\hat{x} = -\nabla^2 \psi(x)^{-1} \nabla f(x)\).

**Example 8:** (Unconstrained) Let \(\Omega^p = \mathbb{R}^n\), \(\partial\psi^p(x^p) = \frac{1}{2} [x^p, \nabla^2 \psi(x^p)] \in \mathbb{R}^n\), \(\nabla^2 \psi(x^p) = \epsilon \mathbf{I}\) where \(\psi^p_c = \epsilon \mathbf{H}^p\). By comparing with NG2 we obtain

\[
\hat{\xi}^p = \beta(x^p - \xi^p), \quad \dot{\xi}^p = \gamma e^{-1}(U^p(x) - \alpha(x^p - \xi^p))
\]

which recovers the one recently introduced in [30]. Equation (23) can be further shown via discretization to recover algorithms such as optimistic gradient-descent/ascent, Polyak’s HB method, among others (see [30]). While (23) is by far the most standard choice for unconstrained action sets, more than one type of dynamics can reside on the same action sets, see [64].

**Example 9:** (Orthant) Let \(\Omega^p = \mathbb{R}_{+}^n\) and consider \(\partial\psi^p(x^p) = \sum_{i=1}^{n} p_i \log(x^p) - x^p_i\) with \(\log(0) = 0\). We can show \(\nabla^2 \psi(x^p) = \epsilon \mathbf{I}\) which occurs \(\psi^p_c = \epsilon \mathbf{H}^p\). Using NG2,

\[
\hat{\xi}^p = \beta(x^p - \xi^p), \quad \dot{\xi}^p = \gamma e^{-1}(U^p(x) - \alpha(x^p - \xi^p)).
\]

The first-order dynamics (\(\alpha = \beta = 0\)) of (24) was recently studied by [56] as the mean-dynamics to a payoff-based dynamics in continuous games with nonnegative orthant action sets. Suppose instead \(\partial\psi^p(x^p) = -\sum_{i=1}^{n} \ln(x^p_i)\) which leads to \(\nabla^2 \psi(x^p) = \epsilon \mathbf{I}\) and

\[
\hat{\xi}^p = \beta(x^p - \xi^p), \quad \dot{\xi}^p = \gamma e^{-1}(U^p(x) - \alpha(x^p - \xi^p)).
\]

We stress however that the log-derivative does not fall under our definition of a regularizer, as \(\text{dom}(\partial\psi^p)\) is a proper subset of \(\Omega^p\). (25) with \(\alpha = \beta = 0\) has been previously studied in [57].

**VII. DISCRETE-TIME SECOND-ORDER DA (DA2) WITH NOISY OBSERVATIONS**

So far we have considered a continuous-time setup whereby each player is able to acquire a partial-gradient \(U^p\) at each time instance. In practical scenarios where the games are played in discrete-time, the acquired pseudogradient information could be corrupted due to a multitude of reasons, such as noisy communication channel. This leads us to consider the so-called “noise-corrupted” pseudogradient scenario (also known as “seminadit” learning). In this setting, each player receives a realization of a so-called noise-corrupted version of the true pseudogradient

\[
\hat{U}^p_{k+1} := U^p(x_k) + \epsilon \mathbf{H}^p_{k+1}
\]

where \(\epsilon \mathbf{H}^p\) is some noise process. A special case of the seminadit scenario is when the payoff is observed as the expectation of the true payoff, i.e., \(U^p(x_k, x_{-k}^p) = \mathbb{E}[U^p(x_k, x_{-k}^p, \epsilon^p)]\), \(U^p\) some random vector, where \(\mathbb{E}\) denotes the expectation operator. Here, \(\hat{U}^p\) is an estimate of the expected partial-gradient \(\nabla^2 x_k^p \mathbb{E}[U^p(x_k, x_{-k}^p, \epsilon^p)]\).

Let us consider the convergence of the discrete-time MD2 with noisy observations, which we refer to as DA2

\[
\begin{align*}
X^p_0 &= C^p_0(Z^p_0) \\
Z^p_{k+1} &= Z^p_k + \gamma \tau_{k+1}(\hat{U}_{k+1} - \alpha(X^p_k - \Xi^p_k)) \\
\Xi^p_{k+1} &= \Xi^p_k + \tau_{k+1}(\beta(X^p_k - \Xi^p_k))
\end{align*}
\]

(26)

where \(X^p, Z^p, \Xi^p\) are the stochastic counter-part of (respectively) \(x^p, z^p, \Xi^p\) at time \(k\), \(\{\tau_k\}_{k \in \mathbb{N}}\), \(\{\gamma\}_{k \in \mathbb{N}}\) denote deterministic sequences of nonincreasing step-sizes, assumed to be common for all players. \(\alpha, \beta, \gamma > 0\) are the auxiliary parameters as from before. When \(\alpha = \beta = 0\), we refer to the resulting expression as the DA with noisy observations

\[
\begin{align*}
X^p_k &= C^p_k(Z^p_k) \\
Z^p_{k+1} &= Z^p_k + \gamma \tau_{k+1}(\hat{U}_{k+1})
\end{align*}
\]

(27)

which coincides with the DA scheme studied in [10] \(\gamma = 1\).

To analyze the convergence behavior of DA2, we impose the following set of regularity assumptions [58, 59].

**Assumption 5:**

\[
\sum_{k \in \mathbb{N}} \tau_k = \infty \sum_{k \in \mathbb{N}} \tau_k^2 < \infty, \lim_{k \to \infty} \tau_k = 0.
\]

(28)

**Assumption 6:** \((L^2\text{-bounded martingale difference noise})\) Assume \(\{\xi_k\}_{k \in \mathbb{N}}, \xi_k = \{(\epsilon^p_k)_{p \in \mathbb{N}}\}\) is a \(L^2\)-bounded marginal difference process adapted to the filtration \(\{\mathcal{F}_k\}_{k \in \mathbb{N}}\): each \(\xi_k\) is a random vector that is measurable with respect to \(\mathcal{F}_k\) for each \(k\), where each \(\mathcal{F}_k\) is the \(\sigma\)-field, i.e., \(\mathcal{F}_k = \sigma(\Xi_0, Z_0, \xi_0, \ldots, \xi_k)\) and \(\mathcal{F}_k \subseteq \mathcal{F}_{k+1}\). In particular, \(\{\xi_k\}_{k \in \mathbb{N}}\) satisfies

\[
\mathbb{E}[|\xi_{k+1}|^2 | \mathcal{F}_k] \leq \sigma^2, \forall k \in \mathbb{N} \text{ a.s.}
\]

(29)

Finally, we impose “global integrability” on MD2, i.e., MD2 has a complete vector field. To do so, we need to convert MD2 into a first-order system by defining: \(\omega^p := (\xi^p, z^p)\), which generates the following stacked system on \(\mathbb{R}^{2n} \cong \mathbb{R}^n \times \mathbb{R}^n\):

\[
\omega^p = \left[\begin{array}{c} \dot{\xi}^p \\ \dot{z}^p \end{array}\right] = \left[\begin{array}{c} \beta(C^p_\xi(z^p) - \xi^p) \\ \beta(C^p_z(z^p) - \alpha(C^p(z^p) - \xi^p)) \end{array}\right]
\]

(30)

or equivalently, \(\omega^p = F^p(\omega)\).

Then we proceed to impose the following assumption on the overall system:

\[
\dot{\omega} = F(\omega), \omega = (\omega^p)_{p \in \mathbb{N}}, F = (F^p)_{p \in \mathbb{N}}.
\]

(31)

**Assumption 8:** (Global integrability) The vector field \(F : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}\) of (31) is continuously globally integrable, that is, for every initial condition \(\xi(0), z(0)\in \mathbb{R}^{2n}\), the unique solution of (31) is defined for all \(t \in \mathbb{R}\).

**Remark 7:** This is satisfied whenever 1) or 2) holds.

1) Suppose \(C^p_\xi\) is bounded on \(\mathbb{R}^{2n}\) and \(U^p\) is bounded continuous locally Lipschitz on \(\text{ran}(C^p_\xi)\) for all \(p \in \mathbb{N}\), then \(F(31)\) is bounded locally Lipschitz and hence continuously globally integrable. This follows from [58].

2) Suppose \(U^p\) is continuous locally Lipschitz and both \(C^p_\xi\) and \(U^p \circ C_\xi\) are sublinear for all \(p\), that is

\[
\lim_{\|z\| \to \infty} \frac{\|C^p_\xi\|_{\mathbb{R}^p}}{\|z\|} < \infty, \lim_{\|z\| \to \infty} \frac{\|U^p \circ C_\xi\|_{\mathbb{R}^p}}{\|z\|} < \infty
\]

(32)

then \(F\) is sublinear and hence continuous and globally integrable. This follows from [60].
Under Assumptions 5–8, DA2 can be shown to track the continuous trajectories generated by MD2 via stochastic approximation arguments [58], [59].

Theorem 4: Let \( \mathcal{G} \) be a concave game and assume that every interior NE is globally VS. Suppose that all players choose strategies according to DA2 and \( C_r = (C^p_r)_{p \in \mathcal{N}} \) is induced by \( \psi^p_r := \epsilon \theta^p_r \), where \( \theta^p_r \) is \( C^2 \) and satisfies Assumption 3(i). Suppose that Assumptions 5, 6, 7, and 8 hold.

Then, \( x_t \) converges to an interior mere VSS of \( \mathcal{G} \) almost surely. The same conclusions hold whenever \( \theta^p_r \) satisfies Assumption 3(ii) and \( \Omega^p_r \) is compact \( \forall p \).

Remark 8: The closest result to ours is [1, Th. 4.7], but for DA (i.e., \( \alpha, \beta = 0 \)). Assumption 5, Assumption 6 are identical to theirs and Assumption 8 encapsulates the Lipschitz continuous requirement for \( U \). The major departure is Assumption 7, which does not hold under certain circumstances, such as convergence toward (boundary) strict VSS in finite games. Hence our result only deals with interior mere VSS in general. Accounting for boundary VSS involves an extension of the inductive shadowing argument in [10] which we leave for future work.

Remark 9: In addition to DA [10], there are several other algorithms that converge in similar settings. The most notable example is the mirror-prox/optimistic MD algorithm, which converges to mere VSS in perfect-gradient feedback setting [18]. However, the authors of [18] noted that convergence in null-coherent saddle point problems (a game with a mere VSS) fails in the presence of noise. This is a key advantage of DA2 over mirror-prox. Another algorithm is the stochastic iterative Tikhonov method of [42], which converges in the semibandit setting. However, the work in [42] requires the game to be strictly monotone, whereas DA2 does not require monotonicity.

VIII. SIMULATIONS

In this section, we consider three illustrative examples. The first example is the RPS game with a nonnegative payoff for ties as in Example 3. We provide convergence behavior of MD and MD2 as well as MDA toward mere and weak VSS. We then provide the convergence behavior of DA2. The second example concerns a wireless power control game previously studied in [34]. We show that DA2 converges in this game under different noise assumptions and study the effect of the number of players. Finally, we provide an example involving a generative adversarial network, which admits a locally mere VSS. We show that DA2 also converges in this game under different noise assumptions. In lieu of exact estimation of the basin of attraction for locally mere VSS, which is difficult, we deal with all such cases through appropriate initialization.

Example 10. (RPS with nonnegative payoff for ties): Consider the RPS game in Example 3. We set to be \( \ell = 0, w = 2 \) and vary the tie payoff parameter \( \zeta \in [1, 2] \) to induce different VS properties on the NE \( x^* = (x^p)_{p \in \mathcal{N}}, x^p = (1/3, 1/3, 1/3) \). We simulate both MD and MD2 at initial conditions \( z(0) = [3 2 1], \zeta(0) = 0, \gamma, \beta, \alpha \) are kept as 1.

We begin by contrasting the continuous-time MD and MD2. For \( \zeta < 1 \), this game is merely monotone, \( x^* \) is globally merely VS, MD2 converges by Theorem 1 while MD diverges (see Fig. 1). For \( \zeta = 1.1 \), the game is 0.1-weakly monotone, but since \( x^* \) is only 0.1-weak VS therefore it can be considered a nearly mere VSS (see Remark 2). In this case, MD2 still converges while MD approaches a heteroclinic orbit (see Fig. 2).

It is useful to also examine the advantage of MD2 over time-averaged MD (MDA). While MDA does converge for \( \zeta = 1 \) toward the mere VSS, when the equilibrium becomes just slightly weak (\( \zeta = 1.1 \)), it no longer converges and instead approaches a Shapley triangle [24] (see Fig. 3). This shows time-averaging is in general not a panacea to nonconvergence.

Next, we perform experiments for the globally merely VS case (\( \zeta = 1 \)) using DA2, under the same initial condition as before. For each of the simulations, we separately perturb the pseudogradients \( U^p \) (9) with zero-mean Gaussian noise with the same variance \( \sigma_\zeta^2 > 0 \) across all players. Fig. 4 shows that DA2 easily converge in the low variance regimes \( \sigma_\zeta^2 = 1 \). As we employ larger variance, e.g., \( \sigma_\zeta^2 = 10 \), the standard step-sizes no longer leads to convergence: a larger \( t_k \) will amplify the additive noise. Instead, we utilize step-size sequences of the form \( 1/k^\gamma \) and search over both \( \ell \in (0, 1), c > 0 \) for optimal sets of parameters. The simulation with tuned step-size is shown in Fig. 5. The strategies \( X = (X^p)_{p \in \mathcal{N}} \) still converge toward the vicinity of the mere VSS.

Example 11. (Wireless power control game with a nonconcave potential): Consider the wireless power control game in [34], where \( \mathcal{N} = \{1, \ldots, N\} \) network users decide on intensities \( x^p \in \mathbb{R} \) of power flow to send over a wireless network.
σ^2 = 10, t_k = 0.38/k^{0.47}, τ_k = 0.13/k^{0.71}.

The payoff function for each user \( p \in \mathcal{N} \) is modeled as

\[
U^p(x^p; x^{-p}) = \log \left( 1 + \frac{a^p \exp(x^p)}{1 + \sum_{p \neq p'} a^p \exp(x^{p'})} \right) - \mathcal{R}^p(x^p)
\]

where, \( a^p \in (0, 1) \) \( \forall p \) and \( \mathcal{R}^p(x^p) = b^p \log(1 + \exp(x^p)) - c^p x^p \) is the cost of user \( p \) for transmission, \( b^p > 0, c^p \geq 0 \).

It can be shown that the game is a potential game with a nonconcave potential function [64].

Consider an example with \( N = 2 \) and problem parameters, \( a = (a^1, a^2) = (1, 1), b = (b^1, b^2) = (4, 4), c = (c^1, c^2) = (3, 3) \). For all players, \( Z^p(0) \) is sampled uniformly from \( [0, 10] \), \( \mathcal{R}^p(0) = 0 \). The NE of this game can be found at \( x^* = (1.8663, 1.8663) \). The Jacobian at the NE is \( J(x^*) = \text{diag}([-4.308 0]) \), hence by Proposition 1, the NE is a locally mere VSS. By Theorem 4, DA2 converges toward this NE. For each of the following experiments, we simulate for \( T = 10^4 \) steps. We set \( \Omega^p \) to be \([-1000, 1000]^2 \) and set \( C^p \) to be the projection operator, \( C^p(Z^p) = \text{argmin}_{y \in \Omega^p} ||\epsilon^{-1} Z^p - y^p||_2^2 \) with \( \epsilon = 1 \). Unless specified otherwise, all parameters \( \gamma, \alpha, \beta \) are kept as 1. All additive noise are zero-mean Gaussian with the same variance \( \sigma_c^2 \).

We start our experiment with a small variance \( \sigma_c^2 = 0.1 \) and set the step-sizes to be \( t_k = 0.39/k^{0.26}, \tau_k = 0.12/k^{0.64} \). Fig. 6 shows that DA2 converges to the NE. We increase the variance to \( \sigma_c^2 = 10 \) and Fig. 7 shows similar observation.

Next, consider an example with \( N = 10 \) players. The game parameters are

\[
\begin{align*}
    a &= (12,4,2.5,20,20,12,3,1,2) \\
    b &= (15,20,20,14,21,20,16,19,17,17) \\
    c &= (13,12,14,8,10,19,10,12,15,1)
\end{align*}
\]

A variance of \( \sigma_c^2 = 1 \) is applied for the following experiments. We employ step-size sequences \( t_k = 0.4/k^{0.03}, \tau_k = 0.78/k^{0.001} \) across all players. The NE of this game is located at \( x^* = (1.87, 0.41, 0.85, 0.28, -0.10, 16, 0.51, 0.54, 2.01, -2.77) \) (rounded to two decimal places) and was confirmed to be a nearly mere VSS (Remark 2). We adjust all the parameters \( \beta, \alpha \) and step-sizes \( c k^{-i}, i \in (0, 1), c > 0 \) on a per-player basis, which resulted in a much closer convergence to the NE as opposed to applying them across all players uniformly (see Fig. 8 and 9).

**Example 12:** (Learning to generate a Gaussian)

Let \( Z \sim \mathcal{N}(\mu, \sigma^2) \) and \( X \sim \mathcal{N}(\mu, \sigma^2) \) be two random variables. We wish to construct a model \( \mathcal{G}_0 : \mathbb{R}^n \rightarrow \mathbb{R}^n, r, m \geq 1 \), with an unknown, parameter \( \theta \) such that \( \mathcal{G}_0(Z) \) recovers the statistics of \( X \). Following [61], \( \mathcal{G}_0 \) can be constructed through solving

\[
\min_{\theta} \max_{w} \mathbb{E}_X[D_w(X)] - \mathbb{E}_Z[D_w(\mathcal{G}_0(Z))]
\]

where, the model \( \mathcal{D}_w : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is parametrized by an unknown continuous parameter \( w \). This problem can be thought of as a game between the owners of the models \( \mathcal{G}_0 \) and \( \mathcal{D}_w \) whereby \( \theta, w \) are their respective strategies. The owner of \( \mathcal{G}_0 \) (or the designer) varies \( \theta \) and uses \( \mathcal{G}_0(Z) \) to estimate the statistics of \( X \) whereas the owner of \( \mathcal{D}_w \) (or the tester) varies \( w \) to incur the largest penalty possible for the discrepancy between \( \mathcal{G}_0(Z) \) and \( X \).

We construct a more complicated variant of the examples in [27] and [61], whereby we assume that the \( \mathcal{G}_0 \) owner wishes to learn the mean and variance of a one-dimensional Gaussian at the same time: let \( Z \sim \mathcal{N}(0, 1) \), and \( X \sim \mathcal{N}(\mu, \sigma^2) \) see Fig. 10, where \( v \in \mathbb{R}, \sigma \in \mathbb{R}_{\geq 0} \) are the mean and standard deviation of a Gaussian distribution. Let \( \mathcal{D}_w(X) = w_1 X^2 + w_2 X, w = (w_1, w_2) \in \mathbb{R} \times \mathbb{R} \) and \( \mathcal{G}_0(Z) = \theta_1 Z + \theta_2, \theta = (\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R} \).
There are several outstanding issues that could be better understood. First, our work did not provide a thorough convergence proof for local VSS, which are dealt with appropriate initialization. Moreover, we did not address convergence toward some boundary point for MD2, as our proofs relied on an interiority condition associated with interior NE (14). One possible approach for dealing with these boundary solutions is to adopt the approach in [12], by restricting ourselves to the context of finite games, where these boundary solutions are quite relevant. Another basic issue is that it is still not entirely clear to us at this time how MD2 relates to other existing continuous-time dynamics for optimization or games. Aside from the cases that we know of, e.g., [30], it is possible that there are other existing algorithms that appear as specific instances of MD2.

Our work opens up an extensive array of directions on the interface between dynamical systems and games. As a starting point, in the continuous-time, an interesting direction would be to explain the difference between MD and MD2 in ZS games through a volume compressibility perspective as in [23] and [25]. We can also try to craft third or even higher order versions of MD based on the technique in [26]. In the discrete-time case, we can further reduce the information, to that of the zero-order information case (or full-bandit feedback). Moreover, we can tackle the case whereby the game itself has parameters that are time-varying. Finally, it is also worthwhile to examine MD2 in a discrete-time setting whereby the mirror map is replaced with a proximal operator.

APPENDIX

The following two lemmas are standard in this area of literature, see [12], [23], and [27]. In particular, Lemma 2 follows directly from the Legendre theorem [52].

**Lemma 2:** Let $\psi^p: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is closed, proper, strictly convex, and finite-valued over $\mathbb{R}^n$;
- $\nabla \psi^p : \text{int}(\text{dom}(\psi^p)) \to \text{int}(\text{dom}(\psi^p))$ is a homeomorphism with inverse mapping $\nabla \psi^p^{-1} = \nabla \psi^p : C^p$;
- $C^p$ is strictly monotone on $\text{int}(\text{dom}(\psi^p))$.

**Lemma 3:** Let $\psi^p: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is closed, proper, convex, and finite-valued over $\mathbb{R}^n$, i.e., $\text{dom}(\psi^p) = \mathbb{R}^n$;
- $\psi^p$ is continuously differentiable and $\nabla \psi^p : C^p$;
- $C^p$ is surjective from $\mathbb{R}^n$ onto $\text{rint}(\text{dom}(\psi^p))$ whenever $\psi^p$ is steep, and onto $\mathbb{R}^n$ whenever $\psi^p$ is nonsteep.

**Proof:** (Proof of Proposition 1) Let $x, x' \in \Omega$ be two arbitrarily chosen strategies. Consider $\bar{\pi} = \theta x + (1 - \theta)x'$, $\theta \in [0, 1]$. Then the proof for all these claims boil down to the equality condition found in [59, Prop. 1.44], in which it can be shown

$$
(x - x')^\top (U(x) - U(x')) = \int_0^1 (x' - x)^\top \frac{\partial U(\bar{\pi})(x' - x)}{\partial \theta} d\theta.
$$

**Let $x' = x^*$, then $$(x - x^*)^\top U(x^*) \leq 0$$ and we have

$$
(x - x^*)^\top U(x) \leq \int_0^1 (x' - x)^\top \frac{\partial U(\bar{\pi})(x' - x)}{\partial \theta} d\theta.
$$

By the definiteness assumptions of $\mathbf{U}(x)$ on $T_{i1}(x)$ for all $x \in \Omega$ and $\bar{\pi} = \theta x + (1 - \theta)x^*$ is contained in $\Omega$ for any $\theta \in [0, 1]$,
(37) implies statements (i), (ii), (iii). Next, suppose the definiteness condition on $J_{0}(x^{*})$ holds for all $y \in T_{0}(x^{*})$. Since $U$ is assumed to be continuously differentiable, $y^* J_{0}(y) x(y)$ is continuous for all $x \in \Omega$. This means for every $\epsilon > 0$ there exists a $\delta(\epsilon) > 0$ such that for all $x \in \Omega, \| x - x^* \| < \delta(\epsilon) \Rightarrow \| y^* J_{0}(y) x(y) - y^* J_{0}(x^*) y \| < \epsilon$. The latter statement implies $\dot{y}^* J_{0}(x) y < \epsilon + y^* J_{0}(x^*) y$. Since this holds for every $\epsilon > 0$, hence $J_{0}(x)$ shares the same definiteness property as $J_{0}(x^*)$ for all $x$ in some ball $D = \{ x \in \Omega \mid \| x - x^* \| < \delta(\epsilon) \}$. Let $\tau = \theta x + (1 - \theta) x^*, \theta \in [0, 1], x \in D$. This implies (37) holds locally on the set of all $\tau$, and ('i'), ('ii'), ('iii') follows. The isolation condition of (i), ('i') are proven in [1, Prop. 2.7].

**Proof:** (Proof of Theorem 1) i) Consider

$$V(\xi, z) = \frac{\alpha}{2\beta} \| (\xi(t) - x^* \|^2 + \gamma^{-1} \sum_{p \in \mathbb{N}} \psi^p_{\xi}(x^*) - z^T x^* + \psi^p_{\xi}(z^*) \tag{38}$$

Here, $V$ is a composite of a quadratic distance term, which measures the progress of $\xi(t)$ toward $x^*$, added onto a function which consists the collection of all the terms associated with the Fenchel–Young inequality [48, p. 88], also known as the Fenchel coupling in [10]. Since $\psi^p$ satisfies Assumption 3(i), $V(\xi, z)$ is continuous, positive definite (due to the Legendre theorem), $V(\xi, z) = 0$ iff $\xi = x^*$ and $z = C_{e}(x^*)$. Taking the time-derivative of $V(\xi, z)$ along the solutions of MD2, using $x = C_{e}(z), \dot{\psi}^p_{\psi^p} = C^p_{\psi^p}$ and $\dot{\gamma} = C(U(x) - \alpha \beta z)$,

$$\dot{V}(\xi, z) = \frac{\alpha}{\beta} (\xi - x^* \|^2 + \gamma^{-1} (C_{e}(z) - x^*)\| z)$$

Substituting in $\dot{\xi} = \beta (x - \xi)$, we have

$$\dot{V}(\xi, z) \leq -\alpha \beta^{-2} \| (\xi - x^* \| ^2 \leq 0 \tag{39}$$

where, we have used $x^*$ is a mere VSS. Observe that $V(\xi, z) = 0$ only if $\xi = 0 \iff x = \xi$. By the Legendre theorem [52], $\psi^p_{\xi}$ is coercive, which implies $V(\xi, z)$ is coercive (radially unbounded) on $\mathbb{R}^n \times \mathbb{R}^d$ and hence all of its sublevel sets are compact. Let $D_{c} = \{ (\xi, z) \in \mathbb{R}^n \times \mathbb{R}^d \mid |V(\xi, z)| \leq c \}$ be a compact sublevel set for some $c > 0$. Consider $E = \{ (\xi, z) \in D_{c} \mid V(\xi, z) = 0 \} = \{ (\xi, z) \in D_{c} \mid \xi = x = C_{e}(z) \}$.

Let $(\xi(t), z(t))$ be some solution starting in $E$, then $\xi = x = C_{e}(z) \Rightarrow \xi = z = J_{C_{e}}(z) \Rightarrow 0 = J_{C_{e}}(z) \dot{z} = z = J_{C_{e}}(z) \dot{z} = 0.$

This means the largest invariant set contained in $E$ is $S = \{ (\xi, z) \in E \mid \dot{z} = 0 \}$. By LaSalle’s invariance principle [42, Th. 3.3], any solution starting from $D_{c}$ converges to $S$ as $t \to \infty$. On $S, \dot{z} = 0 \Rightarrow U(x) = U(C_{e}(z)) = 0$, which means $x = C_{e}(z)$ converges to some interior NE (which we may denote as $x^*$, since $x^*$ is arbitrary), which by assumption is a globally mere VSS. Global convergence follows from coercivity of $V(\xi, z)$.

ii) Suppose $\Omega^p$ is compact for all players $p$. Let $x^p$ be an $\omega$-limit point of $x(t) = C_{e}(z(t))$ of MD2. Suppose that $x^p$ is not a globally mere VSS $x^*$. Take $\Omega_{e}$ be an open ball around $x^p$. By definition, there exists a sequence $\{ x^{(k)} \}_{k \in \mathbb{N}}, x^{(k)} \in \Omega_{e}$, converging toward $x^p$, where $\{ x^{(k)} \}_{k \in \mathbb{N}}$ is an increasing sequence of times. Following the technique in [19], we build an auxiliary sequence in $\Omega_{e}$ and show that $V$ is unbounded below along this sequence, thereby obtaining a contradiction.

First, note that since $\theta^p$ is $\rho$-strongly convex, $\psi^p_{\xi} = \epsilon \theta^p$ is $\epsilon \rho$-strong convex, and hence by the conjugate correspondence theorem [48, Th. 5.26], $\psi^p_{\xi}$ is $e \rho^p$-smooth, i.e., $C^p_{\psi^p} = \dot{\psi}^p_{\psi^p}$ is $(\epsilon \rho^p)^{-1}$-Lipschitz. Let $\tau > 0$ be a time increment, then, by $z(t_k + \tau) - z(t_k) = \int^{t_k + \tau} U(x(s)) - \alpha(x(s) - \xi(s))ds$, we have

$$\| x(t_k + \tau) - x(t_k) \| \leq \| C_{e}(z(t_k + \tau) - C_{e}(z(t_k)) \|$$

The following inequalities follow:

$$\leq (\epsilon \rho) \| (z(t_k + \tau) - z(t_k)) \|$$

and

$$\leq (\epsilon \rho)^{-1} \gamma \| \int^{t_k + \tau} U(x(s)) - \alpha(x(s) - \xi(s))ds \|$$

$$\leq (\epsilon \rho)^{-1} \gamma \| \sum_{p \in \mathbb{N}} \| \psi^p_{\psi^p} \| _{\Omega_{e}} + \alpha \| x(s) - \xi(s) \| _{\Omega_{e}} \| \leq b(\tau) \tag{40}$$

where, $b(\tau)$ is a $\tau$-dependent bound. Let $B_{e}(x(t_k)) = \{ x \in \Omega \mid x - x(t_k) \| \leq b(\tau), x(t_k) \in \Omega_{e} \}$ be a ball around $x(t_k)$. Since $\Omega_{e}$ is open, there exists some $\delta > 0$, such that $B_{e}(x(t_k)) \subseteq \Omega_{e}$ for all $\tau \in [0, \delta]$ hence $x(t_k + \tau) \in \Omega_{e}$.

Integrating (39), we have

$$V(\xi(t), z(t)) = V_0 - \int^{t} \alpha \| x(t) - \xi(t) \| _{\Omega_{e}}^2 U(x) d\tau \tag{41}$$

where $V_0 := V(\xi(0), z(0))$ is some constant. Since $x^*$ is assumed to be a globally mere VSS, we obtain

$$V(\xi(t), z(t)) \leq V_0 - \int^{t} \alpha \| x(t) - \xi(t) \| _{\Omega_{e}}^2 d\tau \tag{42}$$

Expressing this inequality in terms of the sequence $\{ x(t_k) \}_{k \in \mathbb{N}}$ and $\{ \xi(t_k) \}_{k \in \mathbb{N}}, \| x(t) - \xi(t) \| _{\Omega_{e}} \leq \delta \leq 0$, we have

$$\| x(t_k + \tau) - x(t_k) \| \leq V_0 - \int^{t_k} \alpha \| x(t) - \xi(t) \| _{\Omega_{e}}^2 d\tau \tag{43}$$

for some $\delta$. Since $\Omega$ is compact, $x$ is continuous, $\xi$ is a continuous function of $x$, and $\xi$ is not eventually identically equal to $x$ (otherwise, using the rest point condition for MD2 (14) with our assumption that all interior NE are globally mere, we arrive at $x^p = x^*$ hence a contradiction), $\| x - \xi \| _{\Omega_{e}}^2$ achieves a lower-bound on an interval $[t_k, t_k + \delta]$. Let $c = \max \{ l > 0 \mid \| x(t) - \xi(t) \| _{\Omega_{e}} \leq l, t \in [t_k, t_k + \delta], l = 1, \ldots, \}$. This means

$$V(\xi(t_k + \delta), z(t_k + \delta)) \leq V_0 - \int^{t_k + \delta} \alpha c d\tau \leq V_0 - ac\delta k$$

hence $V$ is unbounded below as $k \to \infty$, a contradiction.

**Lemma 4:** Suppose $\theta^p : \mathbb{R}^n \to \mathbb{R} \cup \{ \infty \}$ satisfies Assumption 3(i) or (ii). Let $h = \sum_{p \in \mathbb{N}} \theta^p, x^p \in \text{dom}(\theta^p), x^0 = \text{dom}(\theta^p), x^p \in \partial \theta^p(x^p), \psi^p_{\xi} = \partial \theta^p_{\xi} \forall p$, then

$$D_h(x^*, x) = \epsilon^{-1} \sum_{p \in \mathbb{N}} \psi^p_{\psi^p}(x^p) - \psi^p_{\psi^p}(z^p) - z^T x^p$$

(46)
where, \( x^* = (x^{p*})_{p\in\mathcal{N}} \), \( x = (x^p)_{p\in\mathcal{N}} \).

**Proof (Proof of Lemma 4)** By definition,

\[
D_h(x^*, x) = \sum_{p\in\mathcal{N}} \partial p(x^{p*}) - \partial p(x^p) - \epsilon^1 z^2 \nabla p(x^{p*} - x^p) \tag{47}
\]

\[
= \epsilon^1 - \sum_{p\in\mathcal{N}} \psi_p(x^{p*}) + \psi_p^*(z^p) - z^p \nabla p(x^p) \tag{48}
\]

where, we used \( \psi_p(x^{p*}) = \epsilon^2 \partial p(x^{p*}), z^p \in \partial p \psi_p(x^p) = \epsilon^2 \partial p \psi_p(z^p), \psi_p^*(z^p) = \epsilon^2 \partial p \psi_p^*(z^p) \). We note that \( \psi_p^*(z^p) = \epsilon^2 \partial p \psi_p^*(z^p) \) follows from the conjugate property \( \epsilon^2 \partial p \psi_p^*(z^p) \iff \epsilon^2 \partial p \psi_p^*(z^p) \). [48, p. 93].

**Proof (Proof of Theorem 2) Consider**

\[
V(\xi, \gamma) = \frac{\alpha \gamma}{2} \| \xi - x^* \|_2^2 + \sum_{p\in\mathcal{N}} \psi_p^*(x^{p*}) - z^p \nabla p(x^p) + z^p \psi_p^*(z^p) \tag{49}
\]

Taking the time-derivative of \( V(\xi, \gamma) \) along MD2, and using \( \nabla \psi^*_p = C_p^*, x^p = C_p^*(z^p), \xi = U(x) - \gamma \alpha \beta \xi \), we have

\[
\dot{V}(\xi, \gamma) = \frac{\alpha \gamma}{2} \| \xi - x^* \|_2^2 + \gamma (C_\epsilon(z) - x^*) ^\top \dot{\xi} + (C_\epsilon(z) - x^*) ^\top \dot{\xi} \tag{50}
\]

\[
= -\frac{\alpha^2 \gamma}{\beta^2} \| \xi - x^* \|_2^2 + \frac{\gamma}{\beta} (\dot{\xi} - x^*) ^\top (C_\epsilon(z) - x^*) ^\top \dot{\xi} + (C_\epsilon(z) - x^*) ^\top U(x) \tag{51}
\]

\[
\leq -\frac{\alpha^2 \gamma}{\beta^2} \| \xi - x^* \|_2^2 + \frac{\gamma}{\beta} \| \dot{\xi} \|_2^2 + (x^*) ^\top U(x) \tag{52}
\]

\[
= -\gamma \epsilon \nabla p(x^p) - (x^*) ^\top U(x) \tag{53}
\]

\[
\dot{V}(\xi, \gamma) = \frac{\alpha \gamma}{2} \| \xi - x^* \|_2^2 + \gamma (C_\epsilon(z) - x^*) ^\top \dot{\xi} + (C_\epsilon(z) - x^*) ^\top \dot{\xi} \tag{54}
\]

Carrying on from (50) and using (54), we have

\[
\int_0^T U^p(y^p; x^p) - U^p(x) \frac{d\tau}{\tau} \leq \int_0^T -\dot{Q}^p(\tau) - \frac{\alpha}{\beta^2} \| \dot{\xi} \|_2^2 \frac{d\tau}{\tau} \tag{55}
\]

\[
\leq Q^p(0) - Q^p(t). \tag{56}
\]

Applying Fenchel's inequality [52] to the function \( Q^p \), we obtain

\[
Q^p(t) \geq \dot{Q}^p(y^p) - \frac{\alpha}{\beta^2} \| y^p \|_2^2. \tag{57}
\]

Hence, \( Q^p(0) - Q^p(t) \leq \dot{Q}^p(0) + \dot{Q}^p(y^p) + \frac{\alpha}{\beta^2} \| y^p \|_2^2 \). Where \( Q^p(0) \) is some finite constant [due to Lemma 3(i)]. Therefore, we obtain

\[
\int_0^T U^p(y^p; x^p) - U^p(x) \frac{d\tau}{\tau} \leq Q^p(0) + \dot{Q}^p(y^p) + \frac{\alpha}{\beta^2} \| y^p \|_2^2. \tag{58}
\]

Since \( \Omega^p \) is assumed to be compact, and the numerator is continuous on all of \( \Omega^p \), by Weierstrass theorem [48, Th. 2.12] a maximum is achieved. Taking the limsup yields the result. When \( \alpha = 0 \), this recovers no-regret bound of MD [21].

**Proof (Proof of Proposition 2) By Lemma 3(ii), \( x^p = C_p^*(z^p) = \nabla \psi^*_p(x^p) \), therefore \( x^p \in \partial \psi^*_p(z^p) \). Since \( \psi^*_p \) is the inverse map of \( \partial \psi^*_p \), for all \( x^p \in \text{dom}(\partial \psi^*_p) = \text{rint}(\Omega^p) \)

\[
\partial \psi^*_p(z^p) \nabla \psi^*_p(x^p) = C_p^*.(x^p) \tag{59}
\]

From (59), we can write \( z^p = \nabla \psi^*_p(x^p) + n(x^p) \) where \( n(x^p) \in N_{\Omega^p}(x^p) \). Plugging this expression into (20), we have

\[
\left\{ \dot{\xi} = \beta - \dot{\xi} \right\}
\]

\[
\circ \dot{x} = \gamma \nabla \psi^*_p (\nabla \psi^*_p(x^p) + n(x^p)) (U^p(x) - \alpha (x^p - \xi)). \tag{60}
\]

To reduce (60) from a set of differential inclusions to a set of ODEs, we will show, for all \( x^p \in \text{rint}(\Omega^p) \)

\[
\nabla \psi^*_p (\nabla \psi^*_p(x^p) + n(x^p)) = \nabla \psi^*_p (\nabla \psi^*_p(x^p)). \tag{61}
\]

Indeed, \( x^p \in \partial \psi^*_p(z^p) \iff z^p \in \partial \psi^*_p(x^p) \iff x^p \in \partial \psi^*_p(z^p) \equiv x^p \nabla \psi^*_p(x^p) + n(x^p) \equiv \partial \psi^*_p(x^p) \).

The last equation, \( \nabla \psi^*_p(x^p) + n(x^p) \equiv \partial \psi^*_p(x^p) \). Since \( n(x^p) \in N_{\Omega^p}(x^p) \), therefore \( n(x^p) \equiv \partial \psi^*_p(x^p) \).

By substituting in (60) and using the single-valuedness of \( \partial \psi^*_p \), we obtain

\[
\nabla \psi^*_p (\nabla \psi^*_p(x^p) + n(x^p)) = \nabla \psi^*_p (\nabla \psi^*_p(x^p)). \tag{62}
\]

Taking the Jacobian yields our desired result.

**Proof (Proof of Theorem 4) For simplicity, we assume that all constants associated with MD2 are set to 1. We break up our proof into the following steps.**

1) Show that DA2 is the stochastic approximation of MD2.
Rearranging the $Z_k^p$ update, \[
\frac{Z_{k+1}^p - Z_k^p}{t_{k+1}} = U_{k+1}^p - U_k^p \]
and taking the expectation with respect to the filtration $F_k = \sigma(\Xi_k, Z_0, \ldots, \Delta_k)$ using $E(U_k^p(X_k)|F_k) = U_k^p(X_k)$, $E(\Delta_k^p|F_k) = 0$, we have

$$E\left(\frac{Z_{k+1}^p - Z_k^p}{t_{k+1}}|F_k\right) = U_k^p(X_k) - (X_k^p - Z_k^p).$$

And similarly for $\Xi_k^p$, we have

$$E\left(\frac{\Xi_{k+1}^p - \Xi_k^p}{t_{k+1}}|F_k\right) = X_k^p - \Xi_k^p.$$

which shows that MD2 is the mean ODE of DA2.

2) Next, we need to build interpolated processes for $Z_k$ and $\Xi_k$ and show that the interpolated processes converge to a rest point of MD2. The full construction process is provided in [64].

3) Let $\Xi_k^p: \mathbb{R}_+ \to \mathbb{R}^{n_p}$ and $Z_k^p: \mathbb{R}_+ \to \mathbb{R}^{n_p}$ denote the continuous functions associated with the abovementioned processes and let $Z = \{Z_p^p\}_{p \in \mathbb{N}}$, $\Xi = \{\Xi_p^p\}_{p \in \mathbb{N}}$ be the stacked-vector of all the individual interpolated processes.

4) Next, define the overall process $W: \mathbb{R}_+ \to \mathbb{R}^n$, $W = (\Xi, Z)$. Define the interpolated process $\hat{W}$. By [54 Prop. 4.1 and 4.2], under Assumption 5–8 $\hat{W}$ is an asymptotic pseudotrajectory of the semiflow $\Phi: \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$ induced by $\hat{\omega} = F(\omega)$ (31), that is

$$\lim_{t \to \infty} \sup_{0 < h < T} \|W(t+h) - \Phi(h, W(t))\|_2 = 0$$

for any $T > 0$.

5) By Assumption 7, $\hat{W}$ has compact closure, i.e., is pre-compact. Since we have shown that $\hat{W}$ is a precompact APT of $\Phi$ induced by $\hat{\omega}$, by [54, Th. 5.7(i)] the limit set

$$L(W) = \bigcap_{t \geq 0} \text{cl}(W([t, \infty)))$$

is internally chain transitive, which by [54, Prop. 5.3], $L(W)$ is a compact invariant set.

6) Recall that $V: \mathbb{R}^{n} \to \mathbb{R}$ was used as the Lyapunov function associated with $\hat{\omega} = F(\omega)$ in Theorem 1. Consider the set of critical points of $V$, $\mathcal{E} = \{(\xi, z) \in \mathbb{R}^{2n}|V = 0\} = \{(\xi, z) \in \mathbb{R}^m \times \mathbb{R}^n|\xi = x \in \mathcal{C}\}$.

As in Theorem 1, since $V(\xi', z') < V(\xi, z)$ for all $(\xi', z') \in \mathbb{R}^{2n}\setminus\mathcal{E}$, $(\xi', z') = \Phi(t, (\xi, z))$, and $V(\xi', z') \leq V(\xi, z)$ for all $(\xi, z) \in \mathcal{E}$, $(\xi', z') = \Phi(t, (\xi, z))$, a Lyapunov function for $\mathcal{E}$.

7) Since $V(\xi, z)$ is a Lyapunov function for $\mathcal{E} \subset \mathbb{R}^{2n}$, and $V(\mathcal{E})$ is a constant, $\text{int}(V(\mathcal{E})) = \emptyset$, therefore by [55, Prop. 3.27], $L(W)$ is contained in $\mathcal{E}$. By assumption every NE is a mere VSS, hence every point in $\mathcal{E}$ is an interior mere VSS and $L(W) \subset \mathcal{E}$, therefore $L(W)$ contains a compact subset of the rest points of $\hat{\omega}$.

8) By the definition of a limit set, for any $W(0)$, the interpolated process $\hat{W}(t)$ converges as $t \to \infty$.

9) From our construction of the interpolated process and the diminishing step-size assumption, the convergence of the interpolated processes $\Xi_k$ and $Z_k$ implies the convergence of $\Xi_k$, $Z_k$, respectively.

10) By continuity of $C_r$, it follows that $X_k = C_r(Z_k)$ converges almost surely an interior mere VSS $x^\ast$.

References

[1] A. S. Nemirovsky and D. B. Yudin, *Problem Complexity and Method Efficiency in Optimization Discrete Mathematics*. Hoboken, NJ, USA: Wiley, 1983.

[2] M. Raginsky and J. Bouchrie, “Continuous-time stochastic mirror descent on a network: Variance reduction, consensus, convergence,” in *Proc. IEEE Conf. Decis. Control*, 2012, pp. 6793–6800.

[3] W. Kraiche, A. Bayen, and P. Bartlett, “Accelerated mirror descent in continuous and discrete time;,” in *Proc. Adv. Neural Inf. Process. Syst.*, 2015.

[4] A. Beck and M. Teboulle, “Mirror descent and nonlinear projected subgradient methods for convex optimization,” *Operations Res. Lett.*, vol. 31, no. 3, pp. 167–175, 2003.

[5] T. T. Doan, S. Bose, D. H. Nguyen, and C. L. Beck, “Convergence of the iterates in mirror descent methods,” *IEEE Control Syst. Lett.*, vol. 3, no. 1, pp. 13–19, Jan. 2019.

[6] Y. Yu and B. Aćžmeš, “RLC-circuits-based distributed mirror descent method,” *IEEE Control Syst. Lett.*, vol. 4, no. 3, pp. 548–553, Jul. 2020.

[7] J. C. Duchi, A. Agarwal, and M. J. Wainwright, “Dual averaging for distributed optimization: Convergence analysis and network scaling,” *IEEE Trans. Automat. Control*, vol. 57, no. 3, pp. 592–606, Mar. 2012.

[8] I. Tsidios, S. Lawlor, and M. G. Rabbat, “Push-sum distributed dual averaging for convex optimization,” in *Proc. IEEE Conf. Decis. Control*, 2012, pp. 5453–5458.

[9] Y. Nesterov, “Primal-dual subgradient methods for convex problems,” *Math. Program.*, vol. 120, no. 1, pp. 221–259, 2009.

[10] P. Mertikopoulos and Z. Zhou, “Learning in games with continuous action sets and unknown payoff functions,” *Math. Program.*, vol. 173, no. 1–2, pp. 465–507, 2019.

[11] P. Mertikopoulos and M. Staudigl, “Convergence to Nash equilibrium in continuous games with noisy first-order feedback,” in *Proc. IEEE 56th Conf. Decis. Control*, 2017, pp. 5609–5614.

[12] P. Mertikopoulos and W. Sandholm, “Learning in games via reinforcement and regularization,” *Math. Operations Res.*, vol. 41, no. 4, pp. 1297–1324, 2016.

[13] L. Flekas, E.-V. Vlatakis-Gkaragounis, T. Lianes, P. Mertikopoulos, and G. Pliouras, “No-regret learning and mixed Nash equilibria: They do not mix,” in *Proc. 34th Adv. Neural Inf. Process. Syst.*, 2020.

[14] P. Mertikopoulos, C. Papadimitriou, and G. Pliouras, “Cycles in adversarial regularized learning,” in *Proc. 29th Annu. ACM-SIAM Symp. Discrete Algorithms*, 2018, pp. 2703–2717.

[15] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, vol. 1 & II. New York, NY, USA: Springer-Verlag, 2003.

[16] J. Rosen, “Existence and uniqueness of equilibrium points for concave N-person games,” *Econometrica*, vol. 33, no. 3, pp. 529–534, 1965.

[17] Z. Zhou, P. Mertikopoulos, S. Athey, N. Bambos, P. W. Glynn, and Y. Ye, “Learning in games with lossy feedback,” in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 5140–5150.

[18] P. Mertikopoulos, H. Zenati, B. Lecoutre, C.-S. Foo, V. Chandrasekhar, and G. Pliouras, “Mirror descent in saddle-point problems: Going the extra (gradient) mile,” in *Proc. Int. Conf. Learn. Representations*, 2019, pp. 223–223.

[19] J. Harsanyi, “Oddness of the number of equilibrium points: A new proof,” *Int. J. Game Theory*, vol. 2, no. 1, pp. 235–250, 1973.

[20] J. Hofbauer and W. H. Sandholm, “Stable games and their dynamics,” *J. Econ. Theory*, vol. 144, no. 4, pp. 1685–1693, 2009.

[21] T. Mioton and M.-G. Coco, “On minfy-variational inequalities and evolutionary stable states of generalized monotone games,” *Operations Res. Lett.*, vol. 49, no. 1, pp. 96–100, 2021.

[22] Z. Zhou, P. Mertikopoulos, N. Bambos, P. Glynn, and C. Tomlin, “Countering feedback delays in multi-agent learning,” in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 5776–5783.
[23] A. Giannou, E.-V. Vlatakis-Gkaragkounis, and P. Mertikopoulos, “Survival of the strictest: Stable and unstable equilibria under regularized learning with partial information,” in *Proc. Conf. Learn. Theory*, 2021, vol. 134, pp. 2147–2148.

[24] M. Benaim, J. Hofbauer, and E. Hopkins, “Learning in games with unstable equilibria,” *J. Econ. Theory*, vol. 144, pp. 1694–1709, 2009.

[25] P. Cuckereney, B. Gaujal, and P. Mertikopoulos, “Penalty-regularized dynamics and robust learning procedures in games,” *Math. Operation Res.*, vol. 40, no. 3, pp. 611–635, 2015.

[26] B. Gao and L. Pavel, “On passivity, Reinforcement learning and higher-order learning in monotone agent games,” *IEEE Trans. Autom. Control*, vol. 66, no. 1, pp. 121–136, Jan. 2021.

[27] B. Gao and L. Pavel, “Continuous-time discounted mirror descent dynamics in monotone concave games,” *IEEE Trans. Autom. Control*, vol. 66, no. 11, pp. 5451–5458, Nov. 2021.

[28] A. Krogh and J. A. Hertz, “A simple weight decay can improve generalization,” in *Proc. Adv. Neural Inf. Process. Syst.*, 1992.

[29] B. T. Polyak, “Some methods of speeding up the convergence of iteration methods,” *USSR Comput. Math. Phys.*, vol. 4, no. 5, pp. 1–17, 1964.

[30] D. Gadjov and L. Pavel, “On the exact convergence to Nash equilibrium in monotone regimes under partial-information,” in *Proc. IEEE 59th Conf. Decis. Control*, 2020, pp. 2297–2302.

[31] J. M. Smith and G. R. Price, “The logic of animal conflict,” *Nature*, vol. 246, no. 5427, pp. 15–18, 1973.

[32] J. M. Smith, *Evolution and the Theory of Games*. Cambridge, U.K.: Cambridge Univ. Press, 1982.

[33] A. Kannan and U. V. Shanbhag, “Distributed computation of equilibria in monotone Nash games via iterative regularization techniques,” *SIAM J. Optim.*, vol. 22, no. 4, pp. 1177–1205, 2012.

[34] T. Tataranko, “Stochastic learning in multi-agent optimization: Communication and payoff-based approaches,” *Automatica*, vol. 99, pp. 1–12, 2019.

[35] H. Attouch, Z. Chbani, J. Fadili, and H. Riahi, “First-order optimization methods,” *SIAM J. Control Optim.*, vol. 58, no. 3, pp. 594–609, Mar. 2020.

[36] M. Benaim, J. Hofbauer, and S. Sorin, “Stochastic approximations and differential inclusions,” *SIAM J. Control Optim.*, vol. 44, no. 1, pp. 328–348, 2005.

[37] M. Benaim and M. Faure, “Stochastic approximation, cooperative dynamics and supermodular games,” *Ann. Appl. Probabil.*, vol. 22, no. 5, pp. 2133–2164, 2012.

[38] C. Daskalakis, A. Ilyas, V. Syrgkanis, and H. Zeng, “Training GANs with optimism,” in *Proc. 6th Int. Conf. Learn. Representation*, 2018.

[39] W. Haddad and V. Chellaboina, *Nonlinear Dynamical Syst. Control*. Princeton, NJ, USA:Princeton Univ. Press, 2011.

[40] T. Tataranko and M. Kamgarpour, “Learning Nash equilibria in monotone games,” in *Proc. IEEE Conf. Decis. Control*, 2019, pp. 3104–3109.

[41] B. Gao and L. Pavel, “Second-order mirror descent: convergence in games beyond averaging and discounting,” 2023, arXiv:2111.09982.