LEARNING IN NONATOMIC GAMES, PART I: FINITE ACTION SPACES AND POPULATION GAMES

Saeed Hadikhanloo

Tweag I/O, Paris, France

Rida Laraki

CNRS (Lamsade, University of Dauphine-PSL), Paris, France
University of Liverpool (Computer Science Department), Liverpool, UK

Panayotis Mertikopoulos*

Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG, 38000, Grenoble, France

Sylvain Sorin

Institut de Mathématiques Jussieu-PRG
Sorbonne Université, UPMC, CNRS UMR 7586, Paris, France

Abstract. We examine the long-run behavior of a wide range of dynamics for learning in nonatomic games, in both discrete and continuous time. The class of dynamics under consideration includes fictitious play and its regularized variants, the best reply dynamics (again, possibly regularized), as well as the dynamics of dual averaging / "follow the regularized leader" (which themselves include as special cases the replicator dynamics and Friedman's projection dynamics). Our analysis concerns both the actual trajectory of play and its time-average, and we cover potential and monotone games, as well as games with an evolutionarily stable state (global or otherwise). We focus exclusively on games with finite action spaces; nonatomic games with continuous action spaces are treated in detail in Part II.

1. **Introduction.** The study of evolutionary dynamics in population games has been the mainstay of evolutionary game theory ever since the inception of the field in the mid-1970's. Such dynamics are typically derived from biological models or economic microfoundations that express the net growth rate of a type (or strategy) within a population as a function of the relative frequency of all other types. Thus, owing to the flexibility of this general framework, these considerations have generated an immense body of literature; for an introduction to the topic, we refer the reader to the masterful treatments of [22, 58], and the more recent, panoramic survey of [47, 48].

Our paper takes a complementary approach and aims to examine the long-run behavior of *learning procedures* in nonatomic games. While the topic of learning in games with a *finite* number of players is fairly well-studied – see e.g., the classic

works of [14] and [9] – the same cannot be said for learning in population games. On that account, our goal in this paper is to provide a synthetic treatment for a wide range of dynamics for learning in nonatomic games, unifying along the way a number of existing results in the literature.

We pursue this goal in two parts. In the current paper, we focus exclusively on games with shared, *finite* action spaces; games with *continuous* (and possibly player-dependent) action spaces are significantly more complicated, so their study is deferred to the companion paper [17].

Overview of the paper. To specify the general framework under consideration one has to:

a) Describe the family of games to be studied: These are anonymous nonatomic games and two concrete classes thereof, potential and monotone games; the relevant details are provided in Section 2.

b) Identify the solution concepts involved: We will mainly focus on Nash equilibria and variants of evolutionarily stable states, expressed throughout as solutions of an associated variational inequality; this is also described in Section 2.

c) Define the dynamics under study: This is the content of Section 3, where we begin by recalling the standard fictitious play procedure of [8] and [43], its regularized variants [19], and its continuous-time analogue, the best reply dynamics of [16]. Subsequently, we introduce a general class of dynamics known as *dual averaging* – or *"follow the regularized leader"* – and which contains as special cases the replicator dynamics of [57] and the projection dynamics of [13].

The dynamics we examine evolve in both discrete and continuous time: the continuous-time analysis is presented in Section 4 and comprises the template for the discrete-time analysis that follows in Section 5. In both cases, our analysis builds on previous work and results by [39, 46, 23, 19, 20, 24] and [33, 34]. In this context, some of the new results that we present in our paper can be summarized as follows:

a) We establish a precise link between (regularized) fictitious play and dual averaging: informally, fictitious play best-responds to the average population state whereas dual averaging best-responds to the average population payoff profile.

b) We determine the equilibrium convergence properties of vanishingly regularized fictitious play in continuous time.

c) We provide a range of conditions for the convergence of dual averaging trajectories to evolutionarily stable states, in both discrete and continuous time (and possibly with a variable learning rate).

Finally, in Section 6, we discuss several extensions of these results beyond single-population nonatomic games, and we provide a preview of the issues that arise in nonatomic games with continuous action spaces – the topic of interest of the follow-up paper [17].

2. Preliminaries.

2.1. **Nonatomic games.** A (single-population) *nonatomic game* is defined by the following primitives:

- A *population of players.* This is modeled by the unit interval $\mathcal{I} = [0, 1]$ endowed with the Lebesgue measure μ. The nonatomic aspect means that the behavior of a set of players with zero measure has no impact on the game (see below).

- A set of *pure strategies* $\mathcal{A} = \{1, \ldots, A\}$. Following [49], the players' *strategic behavior* is defined as a measurable function $\chi \colon \mathcal{I} \to \mathcal{A}$, with $\chi(i) \in \mathcal{A}$ denoting the pure strategy of player i. Given a pure strategy $\alpha \in \mathcal{A}$, the inverse image $\chi^{-1}(\alpha) = \{i \in \mathcal{I} : \chi(i) = \alpha\}$ will be called the set of "α-strategists", and the corresponding pushforward measure $x := \chi \sharp \mu \equiv \mu \circ \chi^{-1}$ on $\mathcal{A}$ will be called the *state* of the population. This measure will be the main variable of interest; for convenience, we will view x as an element of the simplex $\mathcal{X} := \Delta(\mathcal{A})$ and we will write $x_\alpha = \mu(\chi^{-1}(\alpha))$ for the *relative frequency* of α-strategists in the population.

- A family of continuous *payoff functions* $u_\alpha \colon \mathcal{X} \to \mathbb{R}$, $\alpha \in \mathcal{A}$. Specifically, $u_\alpha(x)$ denotes the payoff to α-strategists when the state of the population is $x \in \mathcal{X}$. Collectively, we will also write $v(x) = (u_\alpha(x))_{\alpha \in \mathcal{A}}$ for the population's *payoff vector* at state x.

The above definition specifies a *population game* $\mathcal{G} \equiv \mathcal{G}(\mathcal{A}, v)$, cf. [47] and references therein. In this definition, players are *anonymous* in the sense that their payoffs factor through the state of the population and do not depend on specific players choosing specific strategies. For more intricate classes of nonatomic games (possibly involving idiosyncratic components), see [49] and [28].

2.2. **Equilibria, stability, and variational inequalities.** In population games, an equilibrium is defined as a population state $x^* \in \mathcal{X}$ in which almost all players are satisfied with their choice of strategy. Formally, $x^* \in \mathcal{X}$ is an *equilibrium* of $\mathcal{G}$ if

$$u_\alpha(x^*) \geq u_\beta(x^*) \quad \text{for all } \alpha \in \mathrm{supp}(x^*),\ \beta \in \mathcal{A}, \tag{Eq}$$

where $\mathrm{supp}(x^*)$ denotes the support of x^*. Equivalently, this requirement for $x^* \in \mathcal{X}$ can be stated as a *Stampacchia variational inequality* (SVI) of the form

$$\langle v(x^*), x - x^* \rangle \leq 0 \quad \text{for all } x \in \mathcal{X}. \tag{SVI}$$

Equilibria can also be seen as fixed points of the continuous endomorphism $x \mapsto \Pi_\mathcal{X}(x + v(x))$, where $\Pi_\mathcal{X}$ denotes the Euclidean projector on $\mathcal{X}$. By standard fixed point arguments, it follows that the game's set of equilibria is nonempty; we will denote this set as $\mathrm{Eq}(\mathcal{G})$.

Remark. In operator theory and optimization, the direction of (SVI) is typically reversed because optimization problems are usually formulated as *minimization* problems. The utility maximization viewpoint is more common in the literature on population games, so we will maintain this sign convention throughout.

An alternative characterization of equilibria is by means of the *best-response correspondence* $\mathsf{BR} \colon \mathcal{X} \rightrightarrows \mathcal{X}$, defined here as

$$\mathsf{BR}(x) := \operatorname*{arg\,max}_{x' \in \mathcal{X}} \langle v(x), x' \rangle, \tag{1}$$

We then have the following string of equivalences:

$$x^* \in \mathsf{BR}(x^*) \iff \mathrm{supp}(x^*) \subseteq \mathsf{BR}(x^*) \iff x^* \in \mathrm{Eq}(\mathcal{G}) \tag{2}$$

i.e., the equilibria of the game correspond to the fixed points of BR. One can easily show that, under our assumptions, this correspondence has a closed graph and convex compact non-empty values. By Kakutani's fixed point theorem, $\mathrm{Eq}(\mathcal{G})$ – that is, the set of fixed points of BR – is nonempty and compact.

We will also consider the associated *Minty variational inequality* (MVI) for $x^* \in \mathcal{X}$, namely

$$\langle v(x), x - x^* \rangle \leq 0 \quad \text{for all } x \in \mathcal{X}. \tag{MVI}$$

For concreteness, we will write $\mathrm{SVI}(\mathcal{G})$ and $\mathrm{MVI}(\mathcal{G})$ for the set of solutions of (SVI) and (MVI) respectively. In terms of structural properties, $\mathrm{MVI}(\mathcal{G})$ is convex and, under our continuity assumption for v, it is straightforward to verify that $\mathrm{MVI}(\mathcal{G}) \subseteq \mathrm{SVI}(\mathcal{G}) \equiv \mathrm{Eq}(\mathcal{G})$. In this regard, (MVI) represents an "equilibrium refinement" criterion for $\mathcal{G}$ (see also Section 2.3 below).

In more detail, (MVI) is intimately related to the seminal notion of *evolutionary stability* that was introduced by [29] and formed the starting point of evolutionary game theory. To make this link precise, recall that a population state $x^* \in \mathcal{X}$ is an *evolutionarily stable state* of $\mathcal{G}$ [29, 22, 47], if

$$\langle v(\delta x + (1 - \delta)x^*), x - x^* \rangle < 0 \tag{ESS}$$

for all sufficiently small $\delta > 0$ and all states $x \neq x^*$. Analogously, $x^* \in \mathcal{X}$ is a *neutrally stable state* (NSS) of $\mathcal{G}$ if (ESS) holds as a weak inequality, i.e.,

$$\langle v(\delta x + (1 - \delta)x^*), x - x^* \rangle \leq 0 \tag{NSS}$$

for all sufficiently small $\delta > 0$ and all $x \in \mathcal{X}$. Finally, x^* is said to be a *globally evolutionarily stable state* (resp. *globally neutrally stable state*) of $\mathcal{G}$ if (ESS) [resp. (NSS)] holds for all $\delta \in (0, 1]$. In obvious notation, we will write $\mathrm{ESS}(\mathcal{G})$, $\mathrm{NSS}(\mathcal{G})$, $\mathrm{GESS}(\mathcal{G})$ and $\mathrm{GNSS}(\mathcal{G})$ for the respective set of states of $\mathcal{G}$ (by definition $\mathrm{GESS}(\mathcal{G})$ consists of at most a single point).

Characterizations and implications. A concise characterization of these notions of stability was derived by [56] and [21], who showed that $x^* \in \mathrm{ESS}(\mathcal{G})$ if and only if there exists a neighborhood $\mathcal{U}$ of x^* in $\mathcal{X}$ such that

$$\langle v(x), x - x^* \rangle < 0 \quad \text{for all } x \in \mathcal{U} \setminus \{x^*\}, \tag{3}$$

and, furthermore, if we can take $\mathcal{U} = \mathcal{X}$ above, then $\mathrm{GESS}(\mathcal{G}) = \{x^*\}$. For the corresponding neutral versions, we have $x^* \in \mathrm{NSS}(\mathcal{G})$ if and only if (3) holds on $\mathcal{U}$ as a weak inequality, and $x^* \in \mathrm{GNSS}(\mathcal{G})$ if and only if we can take $\mathcal{U} = \mathcal{X}$ in this case. This characterization of $\mathrm{GNSS}(\mathcal{G})$ corresponds precisely to the solution set of (MVI); in particular, we have the following string of inclusions:

$$\begin{array}{ccccc} \mathrm{GESS} & \subseteq & \mathrm{GNSS} & \equiv & \mathrm{MVI} \\ \cap & & \cap & & \cap \\ \mathrm{ESS} & \subseteq & \mathrm{NSS} & \subseteq & \mathrm{Eq} & \equiv & \mathrm{SVI} \end{array} \tag{4}$$

In general, the above inclusions are all one-way; in the next section, we discuss two important cases where some of them become equivalences.

2.3. **Classes of games.** The following classes of games will play an important role in the sequel:

1. *Potential games:* these are games that admit a *potential function*, i.e., a function $P \colon \mathcal{X} \to \mathbb{R}$ such that $v = \nabla P$ on $\mathcal{X}$, or, more generally,

$$P'(x; x' - x) = \langle v(x), x' - x \rangle \quad \text{for all } x, x' \in \mathcal{X}, \tag{Pot}$$

where $P'(x; z) := \lim_{t \to 0^+} [P(x+tz) - P(x)]/t$ denotes the one-sided directional derivative of P at x along z. If $\mathcal{G}$ is a potential game, any local maximum of P is an equilibrium of $\mathcal{G}$, and any *strict* local maximum of P is an ESS of $\mathcal{G}$. For a detailed treatment of potential games in population games extending the initial study of [38], see [46, 47].

2. *Monotone games:* these are games that satisfy the monotonicity condition
$$\langle v(x') - v(x), x' - x \rangle \leq 0 \quad \text{for all } x, x' \in \mathcal{X}. \tag{Mon}$$
This condition implies that every equilibrium of $\mathcal{G}$ is neutrally stable, so we get the following equivalences:
$$\text{MVI}(\mathcal{G}) \equiv \text{GNSS}(\mathcal{G}) = \text{NSS}(\mathcal{G}) = \text{Eq}(\mathcal{G}) \equiv \text{SVI}(\mathcal{G}) \tag{5}$$
In this class, the existence of equilibria relies only on the minmax theorem [37].

Going further, a game is called *strictly monotone* if (Mon) holds as a strict inequality for all $x' \neq x$. In this case, we have the stronger equivalence
$$\text{GESS}(\mathcal{G}) = \text{ESS}(\mathcal{G}) = \text{Eq}(\mathcal{G}), \tag{6}$$
i.e., all inclusions in (4) become two-way. Moreover, $\text{Eq}(\mathcal{G})$ is a singleton which is also the unique GESS of the game.

The intersection of potential and monotone (resp. *strictly* monotone) games occurs when $\mathcal{G}$ admits a *concave* (resp. *strictly* concave) potential. The most important example of such games is the class of nonatomic congestion games that arise in routing and transportation theory, cf. [11].

2.4. **Regularization.** In the sequel, we will often consider "regularized" best responses that are single-valued. These are defined as follows:

Definition 2.1. A *regularizer on* $\mathcal{X}$ is a function $h: \mathbb{R}^{\mathcal{A}} \to \mathbb{R} \cup \{\infty\}$ such that:

1. h is supported on $\mathcal{X}$, i.e., $\text{dom } h = \{x \in \mathbb{R}^{\mathcal{A}} : h(x) < \infty\} = \mathcal{X}$.
2. h is strictly convex and continuous on $\mathcal{X}$.

Given a regularizer on $\mathcal{X}$, we further define:

a) The associated *choice map* $Q_h: \mathbb{R}^{\mathcal{A}} \to \mathcal{X}$ given by
$$Q_h(v) := \arg\max_{x' \in \mathcal{X}} \{\langle v, x' \rangle - h(x')\}. \tag{7a}$$

b) The *regularized best response* to a population state $x \in \mathcal{X}$:
$$\text{BR}_h(x) := \arg\max_{x' \in \mathcal{X}} \{\langle v(x), x' \rangle - h(x')\} = Q_h(v(x)). \tag{7b}$$

Remark 2.1. By our assumptions for h (continuity and strict convexity), Q_h and BR_h are both well-defined and single-valued as correspondences. Moreover, by the strict convexity of h, it follows that the convex conjugate $h^*(y) = \max_{x \in \mathcal{X}} \{\langle y, x \rangle - h(x)\}$ of h is differentiable and satisfies
$$Q_h(v) = \nabla h^*(v) \quad \text{for all } v \in \mathbb{R}^{\mathcal{A}}. \tag{8}$$

Remark 2.2. By construction, BR_h (approximately) best-responds to strategies, while Q_h best-responds to payoff vectors. For this reason, we will sometimes refer to Q_h as a "payoff-based" regularized best response (as opposed to "strategy-based" regularized best response for BR_h).

Moving forward, for part of our analysis, we will also require one (or both) of the regularity conditions below:

1. h is K-*strongly convex* on $\mathcal{X}$, i.e.,

$$h(\lambda x + (1-\lambda)x') \leq \lambda h(x) + (1-\lambda)h(x') - \frac{K}{2}\lambda(1-\lambda)\|x'-x\|^2 \quad \text{(StrCvx)}$$

    for all $x, x' \in \mathcal{X}$ and all $\lambda \in [0,1]$.

2. The subdifferential ∂h of h admits a *continuous selection*; specifically, writing $\mathcal{X}_h \coloneqq \operatorname{dom} \partial h = \{x : \partial h(x) \neq \varnothing\}$ for the "prox-domain" of h, we posit that there exists a continuous map $\widetilde{\nabla} h \colon \mathcal{X}_h \to \mathbb{R}^{\mathcal{A}}$ such that

$$\widetilde{\nabla} h(x) \in \partial h(x) \quad \text{for all } x \in \mathcal{X}_h. \quad \text{(Diff)}$$

By replacing best responses with regularized best responses in the definition of equilibria, we also define the notion of a *regularized equilibrium* as follows:

Definition 2.2. Given a regularizer h on $\mathcal{X}$ and a regularization weight $\varepsilon > 0$, we define an ε-*regularized equilibrium* of $\mathcal{G}$ as a profile $x^* \in \mathcal{X}$ such that $x^* \in \mathsf{BR}_{\varepsilon h}(x^*)$. The set of ε-regularized equilibria of $\mathcal{G}$ will be denoted as $\mathrm{Eq}_{\varepsilon h}(\mathcal{G})$.

Under (Diff), $x^* \in \mathcal{X}$ is an ε-regularized equilibrium if and only if

$$\langle v(x^*) - \varepsilon \widetilde{\nabla} h(x^*), x^* - x \rangle \geq 0 \quad \text{for all } x \in \mathcal{X}. \quad (9)$$

In this regard, ε-regularized equilibria can be seen as Nash equilibria of the "ε-regularized" game $\mathcal{G}_{\varepsilon} \equiv \mathcal{G}_{\varepsilon}(\mathcal{A}, v_{\varepsilon})$ with payoff profile $v_{\varepsilon}(x) \coloneqq v(x) - \varepsilon \widetilde{\nabla} h(x)$; more concisely, $\mathrm{Eq}_{\varepsilon}(\mathcal{G}) = \mathrm{Eq}(\mathcal{G}_{\varepsilon})$.

Equilibria and regularized equilibria are related by the next hemicontinuity property, itself a consequence of the maximum theorem of [6].

Lemma 2.3. *Let $x_{\varepsilon}^* \in \mathrm{Eq}_{\varepsilon}(\mathcal{G})$, $\varepsilon > 0$, be a family of regularized equilibria of $\mathcal{G}$. Then, in the limit $\varepsilon \to 0$, any accumulation point of x_{ε}^* is an equilibrium of $\mathcal{G}$.*

We close this section with a concrete illustration of the above concepts:

Example 2.1 (Logit choice)**.** An important example of regularization corresponds to the entropic regularizer $h(x) = \sum_{\alpha \in \mathcal{A}} x_{\alpha} \log x_{\alpha}$ which, by a standard calculation, leads to the choice map

$$\Lambda(v) = \frac{(e^{v_1}, \ldots, e^{v_A})}{e^{v_1} + \cdots + e^{v_A}}. \quad (10)$$

This map is commonly referred to as the *logit choice map* [30], and the corresponding fixed points $x^* = \Lambda(v(x^*)/\varepsilon)$ are known as *logit equilibria*; for a detailed discussion, see [12, 31, 47], and references therein. It is also well known that the entropic regularizer is 1-strongly convex relative to the L_1-norm, so Λ is 1-Lipschitz continuous, cf. [50]. $\diamond$

3. **Dynamics.** In this section, we introduce a wide range of dynamics for learning in nonatomic games, in both continuous and discrete time. The unifying principle of the dynamics under study is that agents play a best response – regularized or otherwise – to some version of the history of play: the empirical population frequencies, or the players' average gains. For historical reasons, we begin with the discrete-time framework in Section 3.1, and we present the corresponding continuous-time dynamics in Section 3.2.

A note on microfoundations. Throughout the sequel, we will consider models that directly prescribe the state of the population at any given time (X_n or x_t, in discrete or continuous time respectively). This should not be taken to mean that we posit the existence of an exogenous mechanism that somehow *imposes* this population state; instead, this should simply be viewed as a shortcut for the following *endogenous* action selection mechanism (illustrated for concreteness in a discrete-time context):

1. At each stage $n = 1, 2, \ldots$, the players observe some common state variable Z_n that depends only on the state of the population and its history up to stage n (typically Z_n will either be X_n or its historical average $\bar{X}_n = (1/n) \sum_{k=1}^{n} X_k$).
2. Subsequently, players select pure strategies based on the law $X_{n+1} \in f(Z_n) = (f_\alpha(Z_n))_{\alpha \in \mathcal{A}}$ where f is a suitable set-valued mapping defining the players' dynamics (the same throughout the population).
3. Given that the population of players is nonatomic and its distribution at time $n + 1$ follows X_{n+1}, the share of α-strategists at time $n + 1$ will be $X_{\alpha, n+1}$, and the corresponding population state will be X_{n+1}. In turn, X_{n+1} defines the corresponding state variable Z_{n+1} of stage $n + 1$, and the process repeats.

This probabilistic interpretation of population states is standard in the field, so we do not discuss it in more detail here; the interested reader may wish to consult the detailed treatments of [58, 14, 47].

3.1. **Discrete-time processes.** In the discrete-time case, we will consider the evolution of the population state $X_n \in \mathcal{X}$ over a series of epochs $n = 1, 2, \ldots$ (corresponding for example to generations in an evolutionary context). In turn, this population state is determined by a recursive update rule – or *algorithm* – that defines the dynamics in question. Depending on the manner of "best-responding" and the signal that the agents are best-responding to, we have the following dynamics.

Fictitious play. Dating back to [8] and [43], *fictitious play* (FP) posits that agents best-respond to the *empirical frequency* (or *"time average"*) $\bar{X}_n = (1/n) \sum_{k=1}^{n} X_k$ of their population state. Concretely, this means that the population state evolves according to the process

$$X_{n+1} \in \mathsf{BR}(\bar{X}_n) = \arg\max_{x \in \mathcal{X}} \langle v(\bar{X}_n), x \rangle. \tag{FP}$$

In terms of empirical frequencies, we have the recursive dynamics

$$\bar{X}_{n+1} - \bar{X}_n \in \frac{1}{n+1}[\mathsf{BR}(\bar{X}_n) - \bar{X}_n]. \tag{11}$$

In contrast to the sequence of actual population states $X_n \in \mathcal{X}$, $n = 1, 2, \ldots$, the stationary points of the time-averaged process (11) are straightforward to characterize. Indeed, any fixed point x^* of (11) satisfies

$$0 \in \mathsf{BR}(x^*) - x^*, \tag{12}$$

i.e., x^* is stationary under (11) if and only if it is an equilibrium of $\mathcal{G}$.

Regularized fictitious play. A variant of the original fictitious play algorithm is the "regularized" version in which best responses are replaced by *regularized* best responses. In particular, given an underlying regularizer h on $\mathcal{X}$ and a corresponding regularization weight $\varepsilon > 0$, the *regularized fictitious play* (RFP) process is defined as

$$X_{n+1} = \mathsf{BR}_{\varepsilon h}(\bar{X}_n) = \arg\max_{x \in \mathcal{X}} V_{\varepsilon h}(\bar{X}_n, x) \tag{RFP}$$

with

$$V_{\varepsilon h}(x', x) = \langle v(x'), x \rangle - \varepsilon h(x) \tag{13}$$

Then, as before, in terms of empirical frequencies, we have

$$\bar{X}_{n+1} - \bar{X}_n \in \frac{1}{n+1}[\mathsf{BR}_{\varepsilon h}(\bar{X}_n) - \bar{X}_n]. \tag{14}$$

As in the case of (FP), the stationary points of the time-averaged dynamics (RFP) are given by the fixed point equation

$$0 \in \mathsf{BR}_{\varepsilon h}(x^*) - x^*, \tag{15}$$

i.e., they are precisely the ε-regularized equilibria of $\mathcal{G}$ (cf. Definition 2.2).

Remark. In the literature, this process is sometimes referred to as "smooth" or "perturbed" fictitious play, cf. [15, 19] and references therein. This difference in terminology is owed to a different set of assumptions for h, which gives rise to interior-valued choice maps $Q \colon \mathbb{R}^{\mathcal{A}} \to \mathcal{X}$.

Vanishingly regularized fictitious play. When the regularization weight in (RFP) decreases over time, we obtain the *vanishingly regularized fictitious play* (VRFP) dynamics of [3], viz.

$$X_{n+1} = \mathsf{BR}_{\varepsilon_n h}(\bar{X}_n) = \arg\max_{x \in \mathcal{X}} \; V_{\varepsilon_n h}(\bar{X}_n, x) \tag{VRFP}$$

with $\varepsilon_n > 0$ for all n and $\varepsilon_n \to 0$ as $n \to \infty$. In terms of empirical frequencies, we then have

$$\bar{X}_{n+1} - \bar{X}_n \in \frac{1}{n+1}[\mathsf{BR}_{\varepsilon_n h}(\bar{X}_n) - \bar{X}_n]. \tag{16}$$

Because this process is non-autonomous (ε_n depends explicitly on n), it is no longer meaningful to discuss its rest points.

Dual averaging. Dually to the above, instead of best-responding to the aggregate history of *play* (perfectly or approximately), we can also consider the case where agents play a "regularized best response to their aggregate *payoff*". In our context, this gives rise to the *dual averaging* dynamics

$$X_{n+1} = Q_h\left(\eta_n \sum_{k=1}^{n} v(X_k)\right), \tag{17}$$

with Q_h defined in (7a). Then, in iterative form, we get the update rule

$$\begin{aligned} S_n &= S_{n-1} + v(X_n) \\ X_{n+1} &= Q_h(\eta_n S_n) \end{aligned} \tag{DA}$$

where $\eta_n > 0$ is a "learning rate" parameter in the spirit of [42]. In the recursive formulation (DA) of the dynamics, the initialization $S_0 = 0$ is the most common one but, otherwise, it can be arbitrary.

Remark. In the literature on online learning and online optimization, the process (DA) is known as *"follow the regularized leader"* (FTRL) [50, 51]. The terminology "dual averaging" is due to [42] and is more common in the theory of variational inequalities and offline convex optimization. We adopt the latter to highlight the link of our work to variational inequalities.

Links between the dynamics: the case of random matching. To illustrate the relations between the various learning processes discussed above, it will be instructive to consider nonatomic games generated by random matching [58, 47]. In such games, players are randomly matched to play a symmetric two-player game with payoff matrix $A \in \mathbb{R}^{\mathcal{A} \times \mathcal{A}}$, so the population's mean payoff profile is $v(x) = Ax$ for all $x \in \mathcal{X}$. As a result, the dynamics (DA) may be rewritten as

$$X_{n+1} = Q_h\left(\eta_n \sum_{k=1}^{n} v(X_k)\right) = Q_h\big(n\eta_n v(\bar{X}_n)\big) = Q_{h/(n\eta_n)}(v(\bar{X}_n)), \qquad (18)$$

which immediately allows us to recover the variants of fictitious play as follows:

1. For (VRFP): $\eta_n = 1/(n\varepsilon_n)$ for a sequence of weights $\varepsilon_n > 0$, $\lim_{n\to\infty} \varepsilon_n = 0$.
2. For (RFP): $\eta_n = 1/(n\varepsilon)$ for some fixed $\varepsilon > 0$.
3. For (FP): this corresponds to the limiting cases "$n\eta_n = \infty$" or "$h = 0$". [Of course, these parameter choices are not formally allowed in (DA), so this limit is informal.]

Remark. The above is meaningful only when $v(x)$ is linear in x; however, the relation between ε and η will be seen to underlie a large part of the sequel.

3.2. **Continuous-time dynamics.** We now proceed to define the corresponding continuous-time dynamics for each of the processes described above.

Best reply dynamics. The autonomous formulation (11) of fictitious play can be seen as an Euler discretization of the *best reply dynamics* (BRD) of [16], namely

$$\dot{\mu}_t \in \mathsf{BR}(\mu_t) - \mu_t. \qquad \text{(BRD)}$$

In the above, $\mu_t \in \mathcal{X}$ is the continuous-time analogue of the empirical mean process $\bar{X}_n$ in discrete time; we use the notation μ instead of x to highlight this link. Clearly, the rest points of (BRD) are again described by the fixed point equation (12), i.e., μ^* is stationary under (BRD) if and only if it is an equilibrium of $\mathcal{G}$.

Regularized best reply dynamics. Working as above, the *regularized best reply dynamics* (RBRD) are defined as

$$\dot{\mu}_t = \mathsf{BR}_{\varepsilon h}(\mu_t) - \mu_t. \qquad \text{(RBRD)}$$

As in the case of (RFP), the rest points of (RBRD) are characterized by the fixed point equation (15), i.e., μ^* is stationary under (RBRD) if and only if it is an ε-regularized equilibrium of $\mathcal{G}$.

Example 3.1 (The logit dynamics)**.** Building on the logit choice map discussed in Section 2 (cf. Example 2.1), the associated regularized best reply dynamics are known as *logit dynamics*, and are given by

$$\dot{\mu}_t = \Lambda(v(\mu_t)/\varepsilon) - \mu_t. \qquad (19)$$

Vanishing regularized best reply dynamics. If we allow the regularization weight ε to vary in (RBRD), we obtain the non-autonomous dynamics

$$\dot{\mu}_t = \mathsf{BR}_{\varepsilon_t h}(\mu_t) - \mu_t, \qquad \text{(VBRD)}$$

where the "V" indicates again that $\varepsilon_t \to 0$ as $t \to \infty$.

The dual averaging dynamics. Finally, given a regularizer h on $\mathcal{X}$ as above, the dynamics of dual averaging in continuous time can be written as

$$\begin{aligned}\dot{y}_t &= v(x_t)\\ x_t &= Q_h(\eta_t y_t)\end{aligned} \tag{DAD}$$

where $\eta_t \geq 0$ is a time-varying learning parameter, which we assume throughout to be non-increasing. More compactly, the above can also be written as

$$\dot{y}_t = v(Q_h(\eta_t y_t)) \tag{20a}$$

or

$$x_t = Q_h\left(\eta_t \int_0^t v(x_s)\,ds\right), \tag{20b}$$

depending on whether we take a differential- or integral-based viewpoint. Specifically, we note that (20a) is an autonomous ordinary differential equation evolving in the dual, payoff space of the game; by contrast, (20b) is an integral equation stated directly in the primal, strategy space of the game. We also note that, unlike (BRD), (RBRD) and (VBRD), the dynamics (DAD) are stated in terms of the actual population state x_t at time t, not its empirical mean μ_t (which is the driving variable of the previous dynamics).

Example 3.2. If $\eta_t \equiv 1$ and h is the entropic regularizer on $\mathcal{X}$ (cf. Example 2.1), the dynamics (DAD) yield the replicator dynamics of [57], viz.

$$\dot{x}_{\alpha,t} = x_{\alpha,t}[u_\alpha(x_t) - \langle v(x_t), x_t\rangle]. \tag{RD}$$

In words, (RD) indicates that the per capita growth rate of the population of α-strategists is proportional to the difference between the payoff that they experience and the mean population payoff. For a detailed presentation of this derivation in different contexts, see [45, 52, 53, 32] and references therein. $\diamond$

Example 3.3. If $\eta_t \equiv 1$ and $h(x) = (1/2)\sum_{\alpha \in \mathcal{A}} x_\alpha^2$, the integral form (20b) of (DAD) becomes

$$x_t = \Pi\left(\int_0^t v(x_s)\,ds\right) \tag{21}$$

where Π denotes the Euclidean projector on $\mathcal{X}$. The differential form of (21) is more delicate to describe because x_t may enter and exit different faces of $\mathcal{X}$ in perpetuity. However, if we focus on an open interval $\mathcal{T} \subseteq \mathbb{R}$ over which the support of x_t remains constant, it can be shown that (21) follows the projection dynamics of [13], viz. for all $t \in \mathcal{T}$ we have

$$\dot{x}_{\alpha,t} = \begin{cases} u_\alpha(x) - \frac{1}{|\text{supp}(x_t)|}\sum_{\beta\in\text{supp}(x_t)} u_\beta(x) & \text{if } \alpha \in \text{supp}(x_t),\\ 0 & \text{otherwise.}\end{cases} \tag{PD}$$

For the details of this derivation, see [33]. $\diamond$

Remark. For posterity, we note that both regularizers used in the above examples satisfy (StrCvx), (Diff), and the technical condition (Rec) that we describe later in the paper.

Random matching in continuous time. Following [24], the various dynamics presented so far can be linked as follows when v is linear and $\mu_t = \frac{1}{t}\int_0^t x_s\, ds$ for $t > 0$:

$$x_t = \underset{x\in\mathcal{X}}{\arg\max}\left\{\eta_t \int_0^t \langle v(x_s), x\rangle\, ds - h(x)\right\} = \underset{x\in\mathcal{X}}{\arg\max}\{\eta_t\, t\, \langle v(\mu_t), x\rangle - h(x)\}$$
$$= \mathsf{BR}_{[1/(t\eta_t)]h}(\mu_t). \tag{22}$$

It thus follows that the empirical distribution of play μ_t under (DAD) follows the dynamics

$$\dot\mu_t = \frac{\mathsf{BR}_{\varepsilon_t h}(\mu_t) - \mu_t}{t} \tag{23}$$

with $\varepsilon_t = 1/(t\eta_t)$. Hence, after the change of time $t \leftarrow \log t$, we get (RBRD) with a variable regularization parameter $\varepsilon_t = 1/(t\eta_t)$.

4. **Analysis and results in continuous time.** We now proceed to present our results for the class of dynamics under study. We begin with the continuous-time analysis which provides a template for the discrete-time analysis in the next section. We prove that under some assumptions depending on the dynamical system and the class of games under consideration (potential, monotone or strictly monotone), the dynamics converge to the set of equilibria or or – regularized equilibria – of the game.[1]

4.1. **Best reply dynamics.** Our first result concerns the best reply dynamics (BRD). Albeit slightly more general, the results of this subsection and the next essentially follow [19, 20].

Theorem 4.1. *Suppose that one of the following holds:*

(a) *$\mathcal{G}$ is potential.*

(b) *$\mathcal{G}$ is monotone and v is C^1-smooth.*

Then every solution orbit μ_t of (BRD) *converges to* $\mathrm{Eq}(\mathcal{G})$.

Proof. Both cases rely on establishing a suitable Lyapunov function $E(t)$ for (BRD):

1. In the potential case, we simply take $E(t) = P(\mu_t)$.
2. In the monotone case, $E(t) = \mathrm{Gap}(\mu_t)$ where Gap denotes the function

$$\mathrm{Gap}(x) := \max_{x'\in\mathcal{X}}\langle v(x), x' - x\rangle = \langle v(x), \mathsf{BR}(x) - x\rangle \tag{24}$$

Observe that $\mathrm{Gap}(x) \geq 0$ with equality if and only if $x \in \mathrm{Eq}(\mathcal{G})$.

We now proceed on a case-by-case basis:

Potential case. From (Pot) we obtain that

$$\frac{d}{dt}P(\mu_t) = \lim_{\delta\to 0^+}[P(\mu_t + \delta\dot\mu_t) - P(\mu_t)]/\delta = \langle v(\mu_t), \dot\mu_t\rangle \tag{25}$$

Since $\dot\mu_t \in \mathsf{BR}(\mu_t) - \mu_t$, it follows that $\frac{d}{dt}P(\mu_t) = \mathrm{Gap}(\mu_t) \geq 0$ with equality if and only if $\dot\mu_t = 0$. Consequently, from Lyapunov's first method [**?**, see e.g.,]Theorem 2.6.1]HS98 we deduce that μ_t converges to the set of rest points of (BRD), which is $\mathrm{Eq}(\mathcal{G})$.

---

[1]Recall that a family $(x_t)_{t\geq 0}$ (resp. a sequence (X_n), $n = 1, 2, \dots$) converges to a set B if the set of accumulation points $\mathbf{A}[x_t]$ (resp. $\mathbf{A}[x_n]$), as $t \to \infty$ (resp $n \to \infty$), satisfies: $\mathbf{A}[x_t] \subset B$ (resp. $\mathbf{A}[x_n] \subset B$).

Monotone case. Let Jv denote the Jacobian of v. By the envelope theorem we get:

$$\frac{d}{dt}\operatorname{Gap}(\mu_t) = \langle v(\mu_t), -\dot\mu_t\rangle + \dot\mu_t Jv(\mu_t)(\operatorname{BR}(\mu_t) - \mu_t) = -\operatorname{Gap}(\mu_t) + \dot\mu_t Jv(\mu_t)\dot\mu_t \quad (26)$$

As $\mathcal{G}$ is monotone, $(x' - x)Jv(x)(x' - x) \leq 0$ for all $x, x' \in \mathcal{X}$, so $\dot\mu_t Jv(\mu_t)\dot\mu_t \leq 0$. Consequently:

$$\frac{d}{dt}\operatorname{Gap}(\mu_t) \leq -\operatorname{Gap}(\mu_t). \quad (27)$$

Hence $\operatorname{Gap}(\mu_t)$ decreases exponentially to 0, so μ_t converges to $\operatorname{Eq}(\mathcal{G})$. $\qquad\square$

4.2. **Regularized best reply dynamics.** In this section we will assume that the regularizer h satisfies (Diff). In a slight abuse of notation, we will skip the dependence on h and write v_ε for $v_{\varepsilon h} := v - \varepsilon\widetilde\nabla h$, see (9), with $\widetilde\nabla h$ defined as in (Diff). We also write $\operatorname{BR}_\varepsilon(x)$ for $\operatorname{BR}_{\varepsilon h}(x)$ and so on. The reason is that the regularizer h will be fixed throughout while the weight ε will be allowed to vary in the next subsection.

Lemma 4.2. *Under* (Diff), *we have* $\langle v_\varepsilon(x), \operatorname{BR}_\varepsilon(x) - x\rangle \geq 0$ *for all* $x \in \mathcal{X}$, *and equality holds if and only if* $x = \operatorname{BR}_\varepsilon(x)$.

Proof. Let $\phi(z) = \langle v(x), z\rangle - \varepsilon h(z)$. Then ϕ is strictly concave and so $\partial\phi$ defines a strictly monotone operator. Thus, $x' = \arg\max_{z\in\mathcal{X}}\phi(z)$ is a solution of the variational inequality (MVI) associated to $\partial\phi$: for all $d \in \partial\phi(x)$: $\langle d, x' - x\rangle \geq 0$ with equality if and only if $x = x'$. As $v_\varepsilon(x) \in \partial\phi(x)$ and $x' = \operatorname{BR}_\varepsilon(x)$ the result follows. $\qquad\square$

Lemma 4.3. *Suppose that* (Diff) *holds. If* $\mathcal{G}$ *with payoff field* v *is potential (resp. monotone) then the game* $\mathcal{G}_\varepsilon$ *with payoff function* v_ε *is potential (resp. strictly monotone).*

Proof. If $\mathcal{G}$ has a potential P, a potential function of $\mathcal{G}_\varepsilon$ is $P - \varepsilon h$. If $\mathcal{G}$ is a monotone game, the game $\mathcal{G}_\varepsilon$ is strictly monotone because we have $\langle\widetilde\nabla h(x) - \widetilde\nabla h(x'), x - x'\rangle \geq 0$ for every $x, x' \in \mathcal{X}$, with equality if and only if $x = x'$. $\qquad\square$

We now turn to the asymptotic properties of regularized best reply dynamics.

Theorem 4.4. *Suppose that* (Diff) *is satisfied and assume one of the following holds:*

  (a) $\mathcal{G}$ *is potential.*
  (b) $\mathcal{G}$ *is monotone and* v *is* C^1-*smooth.*

Then every solution orbit μ_t *of* (RBRD) *converges to* $\operatorname{Eq}_\varepsilon(\mathcal{G})$.

Remark. Observe that when $\mathcal{G}$ is monotone, $\mathcal{G}_\varepsilon$ is strictly monotone, so $\operatorname{Eq}_\varepsilon(\mathcal{G})$ is a singleton.

Proof. As above, both cases rely on establishing a suitable Lyapunov function $E(t)$ for (RBRD):

  1. In the potential case: $E(t) = P_\varepsilon(\mu_t) := P(\mu_t) - \varepsilon h(\mu_t)$.
  2. In the monotone case:

$$E(t) = \operatorname{Gap}_\varepsilon(\mu_t) = \max_{x'\in\mathcal{X}}\langle v(\mu_t), x' - \mu_t\rangle - \varepsilon(h(x') - h(\mu_t))$$

$$= \max_{x'\in\mathcal{X}} V_\varepsilon(\mu_t, x') - V_\varepsilon(\mu_t, \mu_t) \quad (28)$$

Potential case. There is $z(t) \in \partial h(\mu_t)$ such that – similarly to (25) – we have:

$$\frac{d}{dt} P_\varepsilon(\mu_t) = \lim_{dt \to 0^+} [P_\varepsilon(\mu_t + \dot{\mu}_t dt) - P_\varepsilon(\mu_t)]/dt = \langle v(\mu_t), \dot{\mu}_t \rangle - \varepsilon \langle z(t), \dot{\mu}_t \rangle \qquad (29)$$

As $\dot{\mu}_t = \mathsf{BR}_\varepsilon(\mu_t) - \mu_t$, by Lemma A.1 in the appendix, we have $\langle z(t), \dot{\mu}_t \rangle \leq \langle \widetilde{\nabla} h(\mu_t), \dot{\mu}_t \rangle$. Consequently

$$\frac{d}{dt} P_\varepsilon(\mu_t) \geq \langle v(\mu_t), \dot{\mu}_t \rangle - \varepsilon \langle \widetilde{\nabla} h(\mu_t), \dot{\mu}_t \rangle = \langle v_\varepsilon(\mu_t), \dot{\mu}_t \rangle \qquad (30)$$

By Lemma 4.2 and the above, we deduce that $\frac{d}{dt} P_\varepsilon(\mu_t) \geq 0$ with equality if and only if $\dot{\mu}_t = 0$. Consequently, by a Lyapunov argument, μ_t converges to $\mathrm{Eq}_\varepsilon(\mathcal{G})$.

Monotone case. Using the envelope theorem as in (26) and, just as above, there is $z(t) \in \partial h(\mu_t)$ such that:

$$\frac{d}{dt} \mathrm{Gap}_\varepsilon(\mu_t) = -\langle v(\mu_t) - \varepsilon z(t), \dot{\mu}_t \rangle + \dot{\mu}_t J v(\mu_t) \dot{\mu}_t \leq -\langle v_\varepsilon(\mu_t), \dot{\mu}_t \rangle + \dot{\mu}_t J v(\mu_t) \dot{\mu}_t$$
$$(31)$$

The inequality being a consequence of Lemma A.1 as it implies that $\langle z(t), \dot{\mu}_t \rangle \leq \langle \widetilde{\nabla} h(\mu_t), \dot{\mu}_t \rangle$. Since $\mathcal{G}$ is monotone, $\dot{\mu}_t J v(\mu_t) \dot{\mu}_t \leq 0$. Consequently, as $\dot{\mu}_t = \mathsf{BR}_\varepsilon(\mu_t) - \mu_t$, by Lemma 4.2, $\frac{d}{dt} \mathrm{Gap}_\varepsilon(\mu_t) \leq 0$ with equality if and only if $\dot{\mu}_t = 0$. Hence μ_t converges to $\mathrm{Eq}_\varepsilon(\mathcal{G})$. $\qquad \square$

4.3. **Vanishing regularized best reply dynamics.** A similar technique will be used to obtain convergence to the set of Nash equilibria for vanishing regularized best reply dynamics. The additional difficulty here comes from the fact that the vanishing regularized best reply dynamics is not autonomous and so we should prove by hand the Lyapunov convergence argument. To the best of our knowledge, this is a new result.

Theorem 4.5. *Suppose that* (Diff) *holds and* (VBRD) *is run with a smoothly-varying and strictly decreasing regularization weight* $\varepsilon_t \to 0$. *Assume further that* h *takes values in* $[0, 1]$, *and one of the following holds:*

 (a) $\mathcal{G}$ *is potential.*

 (b) $\mathcal{G}$ *is monotone,* v *is* C^1-*smooth, and* $\widetilde{\nabla} h$ *is bounded.*

Then every solution orbit μ_t *of* (VBRD) *converges to* $\mathrm{Eq}(\mathcal{G})$.

Remark. The assumption that h takes value in $[0, 1]$ is only made for notational convenience and does not incur any loss of generality.

Proof. As before, we treat the potential and monotone cases separately.

Potential case. Let $P_{\varepsilon_t}(\mu_t) = P(\mu_t) - \varepsilon_t h(\mu_t)$. Then, a calculation similar to the above implies that there exists $z(t) \in \partial h(\mu_t)$ such that

$$\frac{d}{dt} P_{\varepsilon_t}(\mu_t) = \langle v(\mu_t) - \varepsilon_t z(t), \mathsf{BR}_{\varepsilon_t}(\mu_t) - \mu_t \rangle - \dot{\varepsilon}_t h(\mu_t) \geq \langle v_{\varepsilon_t}(\mu_t), \mathsf{BR}_{\varepsilon_t}(\mu_t) - \mu_t \rangle - \dot{\varepsilon}_t h(\mu_t) \quad (32)$$

The inequality being a consequence of Lemma A.1 as it implies that $\langle z(t), \mathsf{BR}_{\varepsilon_t}(\mu_t) - \mu_t \rangle \leq \langle \widetilde{\nabla} h(\mu_t), \mathsf{BR}_{\varepsilon_t}(\mu_t) - \mu_t \rangle$. By Lemma 4.2, $\langle v_{\varepsilon_t}(x), \mathsf{BR}_{\varepsilon_t}(x) - x \rangle \geq 0$. Since $-\dot{\varepsilon}_t h(\mu_t) \geq 0$ we deduce that $\frac{d}{dt} P_{\varepsilon_t}(\mu_t) \geq 0$ and consequently $P_{\varepsilon_t}(\mu_t)$ converges. As ε_t goes to zero and $h(\cdot)$ is bounded, $P(\mu_t)$ converges as well. We will deduce from this that μ_t converges to $\mathrm{Eq}(\mathcal{G})$. By contradiction, if a limit point $\mu_0 = \lim_k \mu_{t_k}$ is not in $\mathrm{Eq}(\mathcal{G})$ then:

$$\mathrm{Gap}(\mu_0) = \max_{x' \in \mathcal{X}} \langle v(\mu_0), x' - \mu_0 \rangle = \langle v(\mu_0), \mathsf{BR}(\mu_0) - \mu_0 \rangle = \alpha > 0. \qquad (33)$$

Recall that, by definition, for all t and μ:

$$
\begin{aligned}
\operatorname{Gap}_{\varepsilon_t}(\mu) &= \max_{x' \in \mathcal{X}} \langle v(\mu), x' - \mu \rangle - \varepsilon_t(h(x') - h(\mu)) \\
&= \langle v(\mu), \mathsf{BR}_{\varepsilon_t}(\mu) - \mu \rangle - \varepsilon_t(h(\mathsf{BR}_{\varepsilon_t}(\mu)) - h(\mu)) \\
&= \max_{x' \in \mathcal{X}} V_{\varepsilon_t}(\mu, x') - V_{\varepsilon_t}(\mu, \mu). \tag{34}
\end{aligned}
$$

Also, as h takes values in $[0,1]$, $|V_\varepsilon(x, x') - V_\varepsilon(x, x) - (V_{\varepsilon'}(x, x') - V_{\varepsilon'}(x, x))| \le |\varepsilon - \varepsilon'|$. Thus:

$$
|\operatorname{Gap}_{\varepsilon_t}(\mu) - \operatorname{Gap}(\mu)| \le \varepsilon_t. \tag{35}
$$

As in (25) one has for all k and s:

$$
\frac{d}{dt} P(\mu_{s+t_k}) = \langle v(\mu_{s+t_k}), \mathsf{BR}_{\varepsilon_{s+t_k}}(\mu_{s+t_k}) - \mu_{s+t_k} \rangle \tag{36}
$$

Thus, using (34) then (35) we deduce that:

$$
\frac{d}{dt} P(\mu_{s+t_k}) \ge \operatorname{Gap}_{\varepsilon_{s+t_k}}(\mu_{t_k+s}) - \varepsilon_{s+t_k} \ge \operatorname{Gap}(\mu_{t_k+s}) - 2\varepsilon_{s+t_k}. \tag{37}
$$

As the function $\mu \to \operatorname{Gap}(\mu)$ is uniformly continuous and $t \to \mu_t$ is Lipschitz, recalling that $\operatorname{Gap}(\mu_k) \to \alpha > 0$, we deduce that there is $\delta > 0$ and K_1 such that for $k \ge K_1$ and all $0 \le s \le \delta$ one has:

$$
\operatorname{Gap}(\mu_{t_k+s}) \ge \operatorname{Gap}(\mu_{t_k}) - \alpha/4 \ge 3\alpha/4 \tag{38}
$$

As ε_t decreases to 0 there is K_2 such that for all $k \ge K_2$, $\varepsilon_{s+t_k} \le \alpha/4$. From this, (35) and (38), we obtain that $\frac{d}{dt} P(\mu_{s+t_k}) \ge \alpha/2$ for all $0 \le s \le \delta$, for all $k \ge \max(K_1, K_2)$. Thus $P(\mu_{\delta+t_k}) - P(\mu_{t_k}) = \int_0^\delta \frac{d}{dt} P(\mu_{s+t_k}) ds \ge \delta\alpha/2$, which contradicts the convergence of $P(\mu_t)$.

Monotone case. Using the envelope theorem, and a calculation similar to the above implies that there exists $z(t) \in \partial h(\mu_t)$ such that

$$
\frac{d}{dt} \operatorname{Gap}_{\varepsilon_t}(\mu_t) = -\langle v(\mu_t) - \varepsilon_t z(t), \dot\mu_t \rangle + \dot\mu_t J v(\mu_t) \dot\mu_t + \dot\varepsilon_t(h(\mu_t) - h(BR_{\varepsilon_t}(\mu_t)))
$$
$$
\tag{39a}
$$
$$
\le -\langle v_{\varepsilon_t}(\mu_t), \dot\mu_t \rangle + \dot\mu_t J v(\mu_t) \dot\mu_t + \dot\varepsilon_t(h(\mu_t) - h(BR_{\varepsilon_t}(\mu_t))) \tag{39b}
$$
$$
\le -\langle v_{\varepsilon_t}(\mu_t), \dot\mu_t \rangle + \dot\varepsilon_t(h(\mu_t) - h(BR_{\varepsilon_t}(\mu_t))) \tag{39c}
$$
$$
= -\langle v_{\varepsilon_t}(\mu_t), BR_{\varepsilon_t} - \mu_t \rangle + \dot\varepsilon_t(h(\mu_t) - h(BR_{\varepsilon_t}(\mu_t))). \tag{39d}
$$

The first inequality is a consequence of Lemma A.1, which implies that $\langle z(t), \dot\mu_t \rangle \le \langle \widetilde{\nabla} h(\mu_t), \dot\mu_t \rangle$. The second inequality is a consequence of $\dot\mu_t J v(\mu_t) \dot\mu_t \le 0$ as $\mathcal{G}$ is monotone.

We now proceed to prove that $\frac{d}{dt}(\operatorname{Gap}_{\varepsilon_t}(\mu_t) + \varepsilon_t) < 0$.

Case 1: If $\operatorname{Gap}_{\varepsilon_t}(\mu_t) = 0$ then $\mu_t = BR_{\varepsilon_t}(\mu_t) = \operatorname{Eq}_{\varepsilon_t}(\mathcal{G})$ thus $\dot\mu_t = 0$. Then (39a) implies that $\frac{d}{dt} \operatorname{Gap}_{\varepsilon_t}(\mu_t) = 0$ and thus $\frac{d}{dt}(\operatorname{Gap}_{\varepsilon_t}(\mu_t) + \varepsilon_t) < 0$.

Case 2: If $\operatorname{Gap}_{\varepsilon_t}(\mu_t) > 0$ then since $h(\cdot) \in [0,1]$ and $\dot\varepsilon_t < 0$ we have:

$$
\dot\varepsilon_t(h(\mu_t) - h(BR_{\varepsilon_t}(\mu_t))) \le -\dot\varepsilon_t \tag{40}
$$

and from Lemma 4.2 and (39d) we deduce that $\frac{d}{dt}(\operatorname{Gap}_{\varepsilon_t}(\mu_t) + \varepsilon_t) < 0$.

Consequently, $\mathrm{Gap}_{\varepsilon_t}(\mu_t) + \varepsilon_t$ converges to $\delta \geq 0$ and since ε_t decreases to zero we conclude that $\mathrm{Gap}_{\varepsilon_t}(\mu_t) \to \delta$. If $\delta > 0$, define:

$$\Delta := -\langle v_{\varepsilon_t}(\mu_t), BR_{\varepsilon_t}(\mu_t) - \mu_t\rangle$$

$$= -\langle v(\mu_t), BR_{\varepsilon_t}(\mu_t) - \mu_t\rangle + \varepsilon_t\langle\widetilde{\nabla}h(\mu_t), BR_{\varepsilon_t}(\mu_t) - \mu_t\rangle. \tag{41}$$

From (34) we further get:

$$\Delta = -\mathrm{Gap}_{\varepsilon_t}(\mu_t) + \varepsilon_t(h(\mu_t) - h(BR_{\varepsilon_t}(\mu_t)) + \varepsilon_t\langle\widetilde{\nabla}h(\mu_t), BR_{\varepsilon_t}(\mu_t) - \mu_t\rangle \tag{42}$$

As $\mathrm{Gap}_{\varepsilon_t}(\mu_t) \to \delta$, and by the assumptions, $h(\cdot)$ and $\widetilde{\nabla}h$ are bounded and ε_t decreases to 0, we deduce that for t large enough:

$$\Delta = -\langle v_{\varepsilon_t}(\mu_t), BR_{\varepsilon_t}(\mu_t) - \mu_t\rangle \leq -\delta/2. \tag{43}$$

Consequently, for t large enough we deduce from (39d), (40) and the last inequality that

$$\frac{d}{dt}(\mathrm{Gap}_{\varepsilon_t}(\mu_t) + \varepsilon_t) \leq -\langle v_{\varepsilon_t}(\mu_t), BR_{\varepsilon_t}(\mu_t) - \mu_t\rangle < -\delta/2, \tag{44}$$

which is impossible because $\mathrm{Gap}_{\varepsilon_t}(\mu_t) + \varepsilon_t \geq 0$. Hence $\delta = 0$ and so μ_t converges to $\mathrm{Eq}(\mathcal{G})$. $\qquad\square$

4.4. **Dual averaging dynamics with a constant learning rate.** Before stating our main result we establish some useful lemmas. Define the regret associated to a trajectory x_t and a reference point $x \in \mathcal{X}$ as:

$$\mathrm{Reg}_x(t) = \int_0^t \langle v(x_s), x - x_s\rangle \, ds \tag{45}$$

Lemma 4.6. *Let x_t be a solution trajectory of* (DAD). *Then:*

$$\mathrm{Reg}_x(t) \leq \max h - \min h. \tag{46}$$

This is a particular case of a more general bound due to [27] (see Lemma 4.10 below). The next more compact proof of this simpler bound follows [7] and [54].

Proof. Define for $x \in \mathcal{X}$ and $y \in \mathbb{R}^{\mathcal{A}}$ the Fenchel coupling

$$F_h(x, y) = h(x) + h^*(y) - \langle x, y\rangle. \tag{47}$$

By the Fenchel-Young inequality, $F_h(x, y) \geq 0$. Recall that from (8) and (20b) $x_t = \nabla h^*(y_t) = Q_h(y_t)$ with $y_t = \int_0^t v(x_s)ds$. Then, following [36, Lemma A.2], a simple derivation gives that for all $x \in \mathcal{X}$:

$$\frac{d}{dt}F_h(x, y_t) = \langle v(x_t), \nabla h^*(y_t) - x\rangle = \langle v(x_t), x_t - x\rangle = -\frac{d}{dt}\mathrm{Reg}_x(t), \tag{48}$$

Consequently, following the reasoning of [54], we get

$$\mathrm{Reg}_x(t) \leq F_h(x, 0) = h(x) + h^*(0) \leq \max h + h^*(0) = \max h - \min h \qquad\square$$

Lemma 4.7. *If a solution trajectory $x_t = x^*$ of* (DAD) *is stationary, then $x^* \in \mathrm{Eq}(\mathcal{G})$.*

Proof. From the previous lemma, for all $t > 0$ and $x \in \mathcal{X}$ one has

$$\langle v(x^*), x - x^*\rangle = \frac{\mathrm{Reg}_x(t)}{t} \leq \frac{\max h - \min h}{t}, \tag{49}$$

Letting $t \to \infty$ implies $x^* \in \mathrm{Eq}(\mathcal{G})$, as was to be shown. $\qquad\square$

Lemma 4.8. *If h satisfies* (StrCvx), *then for all $x \in \mathcal{X}$ and all sequences y_n in $\mathbb{R}^{\mathcal{A}}$.*

$$F_h(x, y_n) \to 0 \implies Q_h(y_n) \to x \tag{50}$$

Proof. See Lemma A.2 in the appendix.   □

We also need to assume the so called "reciprocity' condition which requires that for all $x \in \mathcal{X}$ and all sequences y_n in $\mathbb{R}^{\mathcal{A}}$:

$$Q_h(y_n) \to x \implies F_h(x, y_n) \to 0 \qquad \text{(Rec)}$$

Taken together, (StrCvx) and (Rec) imply that the sequence $x_n = Q_h(y_n)$ converges to $x \in \mathcal{X}$ if and only if $F_h(x, y_n) \to 0$. The terminology "reciprocity" is justified by the fact that, under (Rec), the topology induced by the level sets of the Fenchel coupling becomes equivalent to the ordinary topology on $\mathcal{X}$; for a precise statement, see Lemma A.2 in Appendix A. This is a subtle – but important – requirement, which is fortunately satisfied by all the standard regularization choices on $\mathcal{X}$, cf. [35, 36].

We can now state our convergence results for (DAD) when the learning rate is constant.

Theorem 4.9. *Let x_t be a solution trajectory of* (DAD) *with constant learning rate $\eta_t \equiv 1$. Assume further that one of the following holds:*

  (*a*) $\mathcal{G}$ *is potential and h^* is twice differentiable.*

  (*b*) $\mathcal{G}$ *is strictly monotone, and h satisfies* (StrCvx) *and* (Rec).

Then x_t converges to $\mathrm{Eq}(\mathcal{G})$.

Proof. Again, we treat the potential and monotone cases separately.

Potential case. As $y_t = \int_0^t v(x_s) \, ds$ and $x_t = Q_h(y_t) = \nabla h^*(y_t)$ we obtain that:

$$\dot{x}_t = \langle v(x_t), \nabla^2 h^*(y_t) \rangle. \qquad (51)$$

Since h^* is convex, $\nabla^2 h^*$ is semidefinite positive and so:

$$\frac{d}{dt} P(x_t) = \langle v(x_t), \dot{x}_t \rangle = v(x_t) \nabla^2 h^*(y_t) v(x_t) \geq 0, \qquad (52)$$

the inequality being strict whenever $\dot{x}_t \neq 0$. As such, x_t converges to $\mathrm{Eq}(\mathcal{G})$.

Monotone case. As $\mathcal{G}$ is strictly monotone, $\mathrm{Eq}(\mathcal{G})$ is a singleton which coincides with x^* the (necessarily unique) GESS of the game. The calculation in (48) gives:

$$\frac{d}{dt} F_h(x^*, y_t) = \langle v(x_t), x_t - x^* \rangle \qquad (53)$$

As x^* is a GESS $\frac{d}{dt} F_h(x^*, y_t) \leq 0$, with equality if and only if $x_t = Q_h(y_t) = x^*$. Suppose there is $\delta > 0$ such that $\langle v(x_t), x_t - x^* \rangle < -\delta < 0$ for all t large enough. Then, $F_h(x^*, y_t)$ goes to $-\infty$ which is impossible because $F_h(\cdot) \geq 0$. Thus there is $t_k \to \infty$ such that $\langle v(x_{t_k}), x_{t_k} - x^* \rangle \to 0$ and so $x_{t_k} = Q_h(y_{t_k}) \to x^*$. By the reciprocity condition (Rec), we deduce that $F_h(x^*, y_{t_k}) \to 0$. As $\frac{d}{dt} F_h(x^*, y_t) \leq 0$ and $F_h(\cdot) \geq 0$, necessarily $F_h(x^*, y_t)$ converges and by the last argument, its limit is zero. Using the strong convexity of h, Lemma 4.8 implies that $x_t = Q_h(y_t)$ converges to x^*.   □

There are several things to note here. First, the proof of convergence to the unique equilibrium in strictly monotone games (part b of Theorem 4.9) extends to all games with a GESS. Second, the same proof implies that in all games, ESS (if exist) are asymptotically stable under (DAD) with constant learning rate (e.g., the dynamics converges to an ESS if it starts sufficiently close to it). Third, Theorem 4.9

concerns the "day-to-day" evolution of (DAD), not the corresponding empirical mean. Last, the theorem does not apply to monotone games which are not strict.

To complete the picture, the next proposition shows the convergence of the empirical mean to the set of Nash equilibria in all monotone games.

Proposition 4.1. *Let $x_t = Q_h(y_t)$ be a solution trajectory of* (DAD) *with constant learning rate $\eta_t \equiv 1$. If the underlying game is monotone, then $\mu_t = \frac{1}{t} \int_0^t x_s ds$ converges to* $\mathrm{Eq}(\mathcal{G})$.

Proof. By Lemma 4.6 and monotonicity of the game, for all $x \in \mathcal{X}$:

$$\max h - \min h \geq \mathrm{Reg}_x(t) = \int_0^t \langle v(x_s), x - x_s \rangle \, ds \geq \int_0^t \langle v(x), x - x_s \rangle \, ds \quad (54)$$

and hence

$$\frac{\max h - \min h}{t} \geq \langle v(x), x - \mu_t \rangle. \quad (55)$$

Consequently, μ_t converges to $\mathrm{MVI}(\mathcal{G})$ which coincides with $\mathrm{Eq}(\mathcal{G})$ as the game is monotone. $\square$

4.5. **Dual averaging dynamics with a variable learning rate.** This subsection extends the results of the previous section on monotone games in two directions: we consider GESSs instead of (strictly) monotone games, and we treat dynamics with a variable learning rate. Other extensions are also possible, but we do not treat them here.

We begin with a technical lemma on the regret incurred by (DAD):

Lemma 4.10. *If x_t is a solution trajectory of* (DAD) *then the regret is bounded as follows:*

$$\mathrm{Reg}_x(t) \leq \frac{\max h - \min h}{\eta_t}. \quad (56)$$

Consequently, if $t\eta_t \to \infty$ and $x_t = x^*$ for all $t \geq 0$, then $x^* \in \mathrm{Eq}(\mathcal{G})$.

Proof. The bound in the regret follows from [27, Theorem 4.1], and the implication follows as in Lemma 4.7 above. $\square$

Lemma 4.10 allows to extend proposition 4.1 to variable learning rates.

Proposition 4.2. *Let $x_t = Q_h(y_t)$ be a solution trajectory of* (DAD)*. If the underlying game is monotone and $t\eta_t \to \infty$ then $\mu_t = \frac{1}{t} \int_0^t x_s \, ds$ converges to* $\mathrm{Eq}(\mathcal{G})$.

The next result extends part b) of theorem 4.9 to variable learning rates. The proof is a little more complicated.

Theorem 4.11. *Let x_t be a solution trajectory of* (DAD) *run with a smooth, non-increasing learning rate η_t. Suppose further that the underlying game $\mathcal{G}$ admits a GESS x^*. If the regularizer satisfies* (StrCvx) *and* (Rec)*, and the dynamics' learning rate satisfies $\lim_{t \to \infty} t\eta_t = \infty$, then x_t converges to x^*.*

Proof. [27, Eq. 29] prove that

$$\frac{d}{dt} \frac{F_h(x^*, \eta_t y_t)}{\eta_t} \leq \langle v(x_t), x_t - x^* \rangle - \frac{\dot{\eta}}{\eta_t^2}(\max h - \min h). \quad (57)$$

Let $\mathcal{F}_\varepsilon = \{y \in \mathbb{R}^{\mathcal{A}} : F_h(x^*, y) \leq \varepsilon\}$ denote the "Fenchel zone" of magnitude ε with respect to x^*. Then it suffices to show that

For all $\varepsilon > 0$, there exists $T(\varepsilon) > 0$ such that $\eta_t y_t \in \mathcal{F}_\varepsilon$ for all $t \geq T(\varepsilon)$. (P)

Thus, by (50) x_t converges to x^*. The proof of (P) follows from the following two claims:

Claim 1. $\eta_t y_t \in \mathcal{F}_\varepsilon$ *infinitely often.*

Proof. If $\eta_t y_t \notin \mathcal{F}_\varepsilon$, then by (Rec), there exists a neighbourhood $\mathcal{U}(\varepsilon)$ of x^* such that $x_t \notin \mathcal{U}(\varepsilon)$. By definition of GESS, there is $c(\varepsilon) > 0$ such that $\langle v(x_t), x_t - x^* \rangle < -2c(\varepsilon)$. As $\dot{\eta}_t / \eta_t^2 \to 0$ (by de l'Hôpital's rule and the assumption that $t\eta_t \to \infty$), there is $T_1(\varepsilon)$ such that for all $t \geq T_1(\varepsilon)$ we have $\frac{\dot{\eta}_t}{\eta_t^2}(\max h - \min h) < c(\varepsilon)$. Consequently by contradiction to the claim, $\frac{d}{dt}\frac{F_h(x^*, \eta_t y_t)}{\eta_t} < -c(\varepsilon)$ for all t large enough and so $\frac{F_h(x^*, \eta_t y_t)}{\eta_t}$ goes to $-\infty$. Since $F_h(x^*, \eta_t y_t) \geq 0$ this is impossible.   §

Claim 2. *There exists $T_2(\varepsilon) \geq T_1(\varepsilon/2)$ such that if $t \geq T_2(\varepsilon)$ and $\eta_t y_t \in \mathcal{F}_\varepsilon$ then there exists $\tau > 0$ such that $\eta_s y_s \in \mathcal{F}_\varepsilon$ for all $s \in [t, t+\tau]$.*

Proof. We need to consider two cases separately:

Case a). If $\eta_t y_t \in \mathcal{F}_{\varepsilon/2}$ then obviously, there is $\tau > 0$ such that $\eta_s y_s \in \mathcal{F}_\varepsilon$ for all $s \in [t, t+\tau]$.

Case b). If $\eta_t y_t \in \mathcal{F}_\varepsilon$ but $\eta_t y_t \notin \mathcal{F}_{\varepsilon/2}$ then, from claim 1, $\frac{d}{dt}\frac{F_h(x^*, \eta_t y_t)}{\eta_t} < -c(\varepsilon/2) < 0$. Thus there exists $\tau > 0$ such that for all $s \in [t, t+\tau]$: $\frac{F_h(x^*, \eta_s y_s)}{\eta_s} \leq \frac{F_h(x^*, \eta_t y_t)}{\eta_t}$. As $\eta_t y_t \in \mathcal{F}_\varepsilon$, $F_h(x^*, \eta_t y_t) \leq \varepsilon$ and since $\eta_s \leq \eta_t$, we deduce that $F_h(x^*, \eta_s y_s) \leq \varepsilon$. Thus $\eta_s y_s \in \mathcal{F}_\varepsilon$ for all $s \in [t, t+\tau]$.   §

From Claims 1 and 2, there exists $t \geq \max(T_2(\varepsilon), T(\varepsilon))$, such that $\eta_t y_t$ belongs to $\mathcal{F}_\varepsilon$ and remains in it forever.   □

5. **Analysis and results in discrete time.** In this section, we continue with the analysis of the discrete-time dynamics presented in Section 3.1, starting with fictitious play and its variants in Section 5.1 before moving on to the dynamics of dual averaging in Section 5.2.

5.1. **Fictitious play and regularized fictitious play.** The convergence of (FP) and (RFP) can be obtained from the continuous-time analysis of Section 4 using the theory of stochastic approximation [2, 4, 5]. Informally, this theory links the asymptotic behavior of differential inclusions of the form

$$\dot{x}_t \in M(x_t) \tag{58}$$

to the limit sets of discrete-time processes satisfying the basic recursion

$$X_{n+1} - X_n \in \gamma_{n+1} M(X_n), \tag{59}$$

where $M \colon \mathcal{X} \rightrightarrows \mathbb{R}^{\mathcal{A}}$ is an upper semicontinuous, compact-convex-valued correspondence on $\mathcal{X}$, and $\gamma_n > 0$ is a vanishing step-size sequence.

In our specific context, the role of M is to be played by the best-reply correspondence BR for (FP), and its regularized variant $\mathsf{BR}_{\varepsilon h}$ for (RFP). Concretely, we have:

Proposition 5.1. *Let $\bar{X}_n = (1/n)\sum_{k=1}^n X_k$, $n = 1, 2, \ldots$, be the empirical frequency of play under a sequence of population states $X_n \in \mathcal{X}$. Then:*

(i) *If X_n is generated by (FP), $\bar{X}_n$ is an asymptotic pseudotrajectory of (BRD).*

(ii) *If X_n is generated by (RFP), $\bar{X}_n$ is an asymptotic pseudotrajectory of (RBRD).*

Informally, the notion of an asymptotic pseudotrajectory (APT) means that $\bar{X}_n$ asymptotically tracks the pseudo-flow of (BRD)/(RBRD) with arbitrary accuracy over windows of arbitrary length; for a precise definition, see [4, Definitions I and 4.1].The proof of Proposition 5.1 follows immediately from the general theory of [4, Propositions 1.3, 1.4, and Theorem 4.2], so we do not present it here.

In view of this "asymptotic approximation" property, it is plausible to expect that $\bar{X}_n$ exhibits the same asymptotic behavior as the corresponding continuous-time dynamics. Our next result makes this intuition precise in two classes of games:

1. Monotone games, i.e., games that satisfy (Mon).
2. Potential games for which the ε-regularized potential $P_\varepsilon = P - \varepsilon h$, $\varepsilon \geq 0$, satisifies the Sard-type condition:

$$P_\varepsilon(\mathrm{Eq}_\varepsilon(\mathcal{G})) \text{ has empty interior.} \tag{S_ε}$$

Our main results for (FP) and (RFP) may then be stated as follows:

Theorem 5.1. *Suppose that $\mathcal{G}$ satisfies* (Mon) *or* (S_ε) *for $\varepsilon = 0$. Then the empirical frequency of play $\bar{X}_n$ under* (FP) *converges to* $\mathrm{Eq}(\mathcal{G})$.

Theorem 5.2. *Suppose* (Diff) *holds and $\mathcal{G}$ satisfies* (Mon) *or* (S_ε) *for some $\varepsilon > 0$. Then the empirical frequency of play $\bar{X}_n$ under* (RFP) *converges to* $\mathrm{Eq}_\varepsilon(\mathcal{G})$.

Sketch of proof. The proof of Theorems 5.1 and 5.2 follows a similar technical path starting with the fact that $\bar{X}_n$ is an APT of (BRD) / (RBRD) respectively (cf. Proposition 5.1 above). In both cases, Theorem 4.3 of [4] shows that the accumulation points of an APT lie in an internally chain transitive set of the underlying continuous-time dynamics.[2] The claims of Theorem 5.1 may then be proved as follows:

1. For games satisfying (Mon), invoke Theorem 4.1 from Section 4 and Theorem 3.23 and Proposition 3.25 of [4].
2. For games satisfying (S_ε) with $\varepsilon = 0$, invoke Theorem 4.1 from Section 4 and Proposition 3.27 of [4].

Finally, for $\varepsilon > 0$, the proof of Theorem 5.2 is entirely analogous, and simply invokes Theorem 4.4 in lieu of Theorem 4.1 in the above chain of implications.   □

5.2. **Dual averaging.** We now proceed to state and prove our main convergence result for the recursion (DA). To do so, we will require the "reciprocity" condition (Rec) that was introduced for the study of the continuous-time dynamics (DAD), namely

$$Q_h(y_n) \to x \implies F_h(x, y_n) \to 0 \tag{Rec}$$

for all $x \in \mathcal{X}$ and all sequences y_n in $\mathbb{R}^{\mathcal{A}}$. We then have:

Theorem 5.3. *Suppose that* (DA) *is run with a regularizer satisfying* (StrCvx) *and* (Rec), *and a vanishing non-increasing learning rate $\eta_n \searrow 0$ such that*

$$\lim_{n \to \infty} \left(\frac{1}{\eta_n} - \frac{1}{\eta_{n-1}} \right) = 0. \tag{60}$$

Then:

- *If x^* is a GESS of $\mathcal{G}$, the sequence of population states X_n converges to x^*.*

---

[2]An internally chain transitive set is a compact invariant set that contains no proper attractors [10].

- *If x^* is an ESS of $\mathcal{G}$, the same conclusion holds provided that $\max\{\eta_n, \eta_n^{-1} - \eta_{n-1}^{-1}\}$ is small enough and (DA) is initialized sufficiently close to x^*.*

Corollary 5.1. *Suppose that $\mathcal{G}$ is strictly monotone and (DA) is run with a regularizer satisfying (StrCvx) and (Rec), and a learning rate of the form $\eta_n \propto 1/n^p$ for $p \in (0,1)$. Then the sequence of population states induced by (DA) converges to the (necessarily unique) equilibrium of $\mathcal{G}$.*

Our proof relies on the use of a suitable "energy inequality", provided here by a deflated variant of the Fenchel coupling:

Lemma 5.4. *Fix a population state $x \in \mathcal{X}$, and let*

$$E_n := \frac{1}{\eta_n} F_h(x, \eta_n S_n) = \frac{1}{\eta_n}[h(x) - h(X_{n+1}) + \eta_n \langle S_n, X_{n+1} - x \rangle], \qquad (61)$$

with S_n as in (DA). Then, for all $n = 1, 2, \ldots$, we have:

$$E_n \leq E_{n-1} + \langle v(X_n), X_n - x \rangle + [h(x) - \min h]\, \delta_n + \frac{F_h(X_n, \eta_{n-1} S_n)}{\eta_{n-1}}, \qquad (62)$$

where $\delta_n = 1/\eta_n - 1/\eta_{n-1}$. If, in addition, h satisfies (StrCvx), we have

$$E_n \leq E_{n-1} + \langle v(X_n), X_n - x \rangle + [h(x) - \min h]\, \delta_n + \frac{\eta_{n-1}}{2K}\|v(X_n)\|_*^2. \qquad (63)$$

Energy inequalities of the form (62) are standard in the analysis of dual averaging schemes, and they go back at least as far as [41, 42] (who introduced the method). The specific formulation above follows [18, Lemma. C.1], so we omit the proof; for a range of similar computations, see [50, Lemma 2.20], [26, p. 42] and [54].

Now, to move forward with the proof of Theorem 5.3, we will require two complementary arguments. The first is a stability result: we show below that, if X_n is sufficiently close to an evolutionarily stable state x^* and the method's learning rate is "small enough", then X_{n+1} remains close to x^*. To state this precisely, let $\mathcal{F}_\varepsilon = \{y \in \mathbb{R}^{\mathcal{A}} : F_h(x^*, y) \leq \varepsilon\}$ denote the "Fenchel zone" of magnitude ε with respect to x^*, and let $\mathcal{D}_\varepsilon = Q_h(\mathcal{F}_\varepsilon) = \{x = Q_h(y) : y \in \mathcal{F}_\varepsilon\}$ denote the image of $\mathcal{F}_\varepsilon$ under Q_h. We then have:

Lemma 5.5. *Let $x^* \in \mathrm{ESS}(\mathcal{G})$ and assume that (StrCvx) and (Rec) hold. Then, for all sufficiently small $\varepsilon > 0$, we have the following implication (valid for all $n = 1, 2, \ldots$): if $X_n \in \mathcal{D}_\varepsilon$ and $\max\{\eta_n, \delta_n\}$ is small enough, we also have $X_{n+1} \in \mathcal{D}_\varepsilon$.*

The second component of the proof of Theorem 5.3 is a subsequence extraction argument: if the iterates of (DA) lie in a neighborhood of x^* where (ESS) holds, there exists a subsequence of X_n that converges to x^*.

Lemma 5.6. *Let $x^* \in \mathrm{ESS}(\mathcal{G})$, let $\mathcal{U}$ be the neighborhood of validity of (ESS), and assume further that (StrCvx) and (60) hold. If $X_n \in \mathcal{U}$ for all sufficiently large n, we have $\liminf_{n \to \infty} \|X_n - x^*\| = 0$.*

We prove these two ancillary results below.

Proof of Lemma 5.5. Let $\mathcal{U}$ be the neighborhood whose existence is guaranteed by (ESS), i.e., $\langle v(x), x - x^* \rangle < 0$ for all $x \in \mathcal{U}$, $x \neq x^*$. By the relation between the Fenchel coupling and the ordinary norm topology on $\mathcal{X}$ (cf. Lemma A.2 in Appendix A), it follows that $\mathcal{D}_\varepsilon \subseteq \mathcal{U}$ if ε is taken small enough; we will assume this to be the case throughout the rest of this proof. Moreover, by (Rec) and the

continuity of u, there exists some $c \equiv c(\varepsilon)$ such that $\langle v(x), x - x^* \rangle \leq -c < 0$ for all $x \in \mathcal{D}_\varepsilon \setminus \mathcal{D}_{\varepsilon/2}$.

With all this in hand, assume that $X_n \in \mathcal{D}_\varepsilon$; we will show that $X_{n+1} \in \mathcal{D}_\varepsilon$ provided that $\max\{\eta_n, \delta_n\}$ is small enough. To do so, let $F_n = F_h(x^*, \eta_n S_n)$; then, by substituting $x \leftarrow x^*$ in the template inequality (63) of Lemma 5.4, we obtain:

Case 1. If $X_n \in \mathcal{D}_\varepsilon \setminus \mathcal{D}_{\varepsilon/2}$, then

$$F_n \leq F_{n-1} - \eta_n \left[c - \delta_n H - \eta_{n-1} \frac{G^2}{2K} \right], \tag{64}$$

where $H = \max h - \min h$ and $G = \max_{x \in \mathcal{X}} \|v(x)\|_*$. Hence, if $\delta_n \leq c/(2H)$ and $\eta_{n-1} \leq cK/G^2$, we conclude that $F_n \leq F_{n-1}$.

Case 2. If $X_n \in \mathcal{D}_{\varepsilon/2}$, then

$$F_n \leq \frac{\varepsilon}{2} + \delta_n H + \eta_{n-1} \frac{G^2}{2K}, \tag{65}$$

so $F_n \leq \varepsilon$ provided that $\delta_n < \varepsilon/(4H)$ and $\eta_{n-1} \leq \varepsilon K/(2G^2)$.

In both cases, we conclude that $X_{n+1} = Q_h(\eta_n S_n) \in \mathcal{D}_\varepsilon$, as claimed. ∎

Proof of Lemma 5.6. Assume to the contrary that $\liminf_{n \to \infty} \|X_n - x^*\| > 0$. Then, by assumption, there exist $\varepsilon > 0$ and $n_0 \equiv n_0(\varepsilon)$ such that $X_n \in \mathcal{U}$ and $\|X_n - x^*\| \geq \varepsilon$ for all $n \geq n_0$. Since $x^* \in \mathrm{ESS}(\mathcal{G})$, there exists a positive constant $c > 0$ such that $\langle v(X_n), X_n - x^* \rangle \leq -c < 0$ for all $n \geq n_0$. With this in mind, substituting $x \leftarrow x^*$ in the template inequality (63) and telescoping yields

$$E_n \leq E_{n_0} - c(n - n_0) + \left(\frac{1}{\eta_n} - \frac{1}{\eta_{n_0}} \right) H + \frac{G^2}{2K} \sum_{k=n_0}^{n} \eta_k$$

$$\leq E_{n_0} + cn_0 - H/n_0 - n \left[c - \frac{1}{n\eta_n} - \frac{1}{n} \sum_{k=n_0}^{n} \eta_k \right]. \tag{66}$$

To proceed, note that (60) gives

$$\lim_{n \to \infty} \frac{(1/\eta_n - 1/\eta_{n-1})}{n - (n-1)} = \lim_{n \to \infty} \left(\frac{1}{\eta_n} - \frac{1}{\eta_{n-1}} \right) = \lim_{n \to \infty} \delta_n = 0, \tag{67}$$

so $\lim_{n \to \infty} 1/(n\eta_n) = 0$ by the Stolz–Cesàro theorem. By (60), we also have $\lim_{n \to \infty} (1/n) \sum_{k=n_0}^{n} \eta_k = 0$. Hence, by letting $n \to \infty$ in (66), we get $\lim_{n \to \infty} E_n = -\infty$, a contradiction which completes our proof. ∎

The proof of our main convergence results follows by a tandem application of the above.

Proof of Theorem 5.3. We consider the two cases separately.

Case 1: $x^* \in \mathrm{GESS}(\mathcal{G})$. Let $\varepsilon > 0$ be arbitrary. By (60), it follows that $\lim_{n \to \infty} \eta_n = \lim_{n \to \infty} \delta_n = 0$, and, by Lemma 5.6, there exists a subsequence X_{n_k} of X_n converging to x^*. This means that, in the long run, η_n and δ_n become arbitrarily small, and X_n comes arbitrarily close to x^*. Hence, by applying Lemma 5.5 inductively, we conclude that $X_n \in \mathcal{D}_\varepsilon$ for all sufficiently large n. Since $\varepsilon > 0$ has been chosen arbitrarily, this implies that X_n converges to x^*, as claimed.

Case 2: $x^* \in \mathrm{ESS}(\mathcal{G})$. Let $\mathcal{U}$ be the neighborhood whose existence is guaranteed by the evolutionary stability condition (3) for x^*. Then, by Lemma 5.5, if (DA) is initialized sufficiently close to x^* and $\max\{\eta_n, \delta_n\}$ is sufficiently small, we will have $X_n \in \mathcal{U}$ for all n. Hence, by Lemma 5.6, we conclude that $\liminf_{n\to\infty} \|X_n - x^*\| = 0$. Our proof then follows by the same inductive application of Lemma 5.5 as in the case of a GESS above. $\qquad\square$

Time averages under dual averaging. We close this section with a result on the empirical frequency of play under (DA). Indeed, telescoping (62) trivially yields the bound:

$$\sum_{k=1}^{n} \langle v(X_k), x - X_k \rangle \leq \frac{\max h - \min h}{\eta_n} + \frac{1}{2K} \sum_{k=1}^{n} \eta_{k-1} \|v(X_k)\|_*^2. \tag{68}$$

This regret-type bound echoes standard results in the literature – see e.g., [42, Theorem 1] – and leads to the following time-average convergence result:

Proposition 5.2. *Suppose that* (DA) *is run with a regularizer satisfying* (StrCvx) *and a learning rate of the form* $\eta_n \propto 1/\sqrt{n}$. *If* $\mathcal{G}$ *is monotone, the empirical frequency of play* $\bar{X}_n$ *under* (DA) *converges to* $\mathrm{Eq}(\mathcal{G})$.

The proof of Proposition 5.2 mimics that of Proposition 4.1 so we omit it. We only note that, in contrast to Theorem 5.3, Proposition 5.2 does not require strict monotonicity or reciprocity; however, it concerns the cruder, average frequency of play, so it provides a considerably weaker guarantee in this regard.

6. **Extensions and concluding remarks.** In this concluding section, we provide a series of extensions and remarks that would have otherwise disrupted the flow of the rest of our paper.

6.1. **Positive correlation.** [46] introduced the so-called "*positive correlation*" condition

$$\dot{x}_t \neq 0 \implies \langle v(x_t), \dot{x}_t \rangle > 0. \tag{PC}$$

In the case of potential games, (PC) implies that

$$\frac{d}{dt} P(x_t) = \langle v(x_t), \dot{x}_t \rangle > 0 \tag{69}$$

whenever $\dot{x}_t \neq 0$. This in turn implies the convergence of the trajectories to the set of rest points of the dynamics [47, Theorem 7.1.2]. Some of our results in Section 4 for potential games can also be obtained by establishing (PC) for the dynamics under study.

6.2. **No-regret dynamics.** Our second remark concerns continuous- or discrete-time processes for which the average regret vanishes, i.e., for all $x \in \mathcal{X}$ we have:

$$\limsup_{t\to\infty} \frac{1}{t} \int_0^t \langle v(x_s), x - x_s \rangle \, ds \leq 0 \quad \text{in continuous time} \tag{70a}$$

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \langle v(X_k), x - X_k \rangle \leq 0 \quad \text{in discrete time} \tag{70b}$$

As we saw in Propositions 4.1 and 5.2, the dynamics of dual averaging have the no-regret property (70) in both continuous and discrete time. Looking more closely at the proofs of Propositions 4.1 and 5.2, we conclude that, in monotone games, the empirical frequency of play under any dynamical process that satisfies (70) converges

to the game's set of equilibria. This property forms the basis of many algorithms and dynamics designed to solve variational inequalities and equilibrium problems [40, 42, 25]; the above provides a game-theoretic interpretation of this property in the context of population games.

6.3. **Extensions.** We also note that our results can be extended to the following settings:

1. If $\mathcal{X}$ is a convex compact subset of $\mathbb{R}^{\mathcal{A}}$ and $v \colon \mathcal{X} \to \mathbb{R}^{\mathcal{A}}$ is a vector field on $\mathcal{X}$, we maintain the same results for convergence to solutions of (SVI) and (MVI); however, the population game interpretation disappears.

2. This interpretation is recovered in multi-population games: if there are several player populations indexed by $i = 1, \ldots, N$, and if $\mathcal{A}_i$ denotes the set of pure strategies of the i-th population and $u_{i\alpha} \colon \prod_{i=1}^N \Delta(\mathcal{A}_i) \to \mathbb{R}$ is the payoff function of α-strategists in the i-th population, it suffices to set $\mathcal{X} = \prod_i \Delta(\mathcal{A}_i)$ and $v(x) = (u_{i\alpha_i}(x))_{\alpha_i \in \mathcal{A}_i, i \in \mathcal{I}}$.

3. Our theory can also be extended to normal form games with a finite number of players and smooth concave payoff functions g_i, $i = 1, \ldots, N$. In this case, the components of the corresponding vector field are $G_{i\alpha}(x) = \frac{\partial g_i}{\partial x_{i\alpha}}$, and the notion of evolutionary stability boils down to that of *variational stability*; for a detailed treatment, see [55] and [36].

6.4. **Links with variational inequalities and further extensions.** Taking a more general point of view that focuses on the vector field v, we can also consider the following framework that brings to the forefront the associated variational inequalities. Let $\mathcal{X}_i$, $i = 1, \ldots, N$ be a convex compact subset of a topological vector space $\mathcal{V}_i$ with dual $\mathcal{V}_i^*$, write $\mathcal{V} = \prod_i \mathcal{V}_i$ and $\mathcal{X} = \prod_i \mathcal{X}_i$, and let $v_i \colon \mathcal{X} \to \mathcal{V}_i^*$ be a collection of continuous maps, $i = 1, \ldots, N$. Finally, set $v(x) = (v_1(x), \ldots, v_N(x))$ and, as usual, write

$$\langle v(x), z \rangle = \sum_{i=1}^N \langle v_i(x), z_i \rangle \tag{71}$$

for the dual pairing between $v(x) \in \mathcal{V}^* = \prod_i \mathcal{V}_i^*$ and $z = (z_1, \ldots, z_N) \in \mathcal{V} \coloneqq \prod_i \mathcal{V}_i$.

We may then define $\mathrm{SVI}(\mathcal{X}, v)$ as the set of $x^* \in \mathcal{X}$ such that

$$\langle v(x^*), x - x^* \rangle \leq 0, \quad \text{for all } x \in \mathcal{X} \tag{72}$$

and $\mathrm{MVI}(\mathcal{X}, v)$ as the set of $x^* \in \mathcal{X}$ such that

$$\langle v(x), x - x^* \rangle \leq 0 \quad \text{for all } x \in \mathcal{X}. \tag{73}$$

The counterparts for ESS and GESS may also be defined similarly (though the link with population games is no longer present).

Example 6.1. Consider a strategic game with N players, compact metric strategy spaces $\mathcal{A}_i \subseteq \mathbb{R}^d$ and continuous payoff functions $u_i \colon \mathcal{A} \to \mathbb{R}$, $i = 1, \ldots, N$, with $\mathcal{A} \coloneqq \prod_j \mathcal{A}_j$. Introduce the mixed extension of the game as usual: for a probability distribution (mixed strategy profile) $\chi = (\chi_1, \ldots, \chi_N) \in \mathcal{X}$ with $\chi_i \in \mathcal{X}_i \coloneqq \Delta(\mathcal{A}_i)$, let u_i be extended as

$$u_i(\chi) = \int_{\mathcal{A}} u_i(\alpha_1, \ldots, \alpha_N) \, d\chi_1(\alpha_1) \cdots d\chi_N(\alpha_N). \tag{74}$$

Then, introduce $v_i(\chi) = u_i(\cdot, \chi_{-i})$ so that

$$\langle v_i(\chi), \chi_i' \rangle = v_i(\chi_i'; \chi_{-i}) \tag{75}$$

for all $\chi_i' \in \mathcal{X}_i = \Delta(\mathcal{A}_i)$. In this way, the equilibrium condition for $\chi^* \in \mathcal{X}$ becomes

$$\langle v(\chi^*), \chi - \chi^* \rangle = \sum_{i=1}^{N} \langle v_i(\chi^*), \chi_i - \chi_i^* \rangle \leq 0 \quad \text{for all } \chi \in \mathcal{X}, \tag{76}$$

which is again an instance of $\mathrm{SVI}(\mathcal{X}, v)$. $\diamond$

In this setting, v is monotone (dissipative) if

$$\langle v(x) - v(x'), x - x' \rangle \leq 0 \quad \text{for all } x, x' \in \mathcal{X} \tag{77}$$

As in the case of Section 2, we have $\mathrm{SVI}(\mathcal{X}, v) = \mathrm{MVI}(\mathcal{X}, v)$ if v is monotone; furthermore, mutatis mutandis, all the properties presented in Section 2.3 extend to this general setup.

The best-response map associated to v is defined again as

$$\mathsf{BR}(x) = \arg\max_{x' \in \mathcal{X}} \langle v(x), x' \rangle. \tag{78}$$

In addition to monotonicity, we will also now make the "*unique best response*" assumption

$$\mathsf{BR}(x) \text{ is a singleton for all } x \in \mathcal{X}. \tag{U}$$

Proposition 6.1. *If v is monotone and* (U) *holds, the set $E \coloneqq \mathrm{MVI}(\mathcal{X}, v) = \mathrm{SVI}(\mathcal{X}, v)$ is a singleton.*

Proof. If $x, x' \in E$, we have $\langle v(x), x' - x \rangle \leq 0$ and $\langle v(x'), x - x' \rangle \leq 0$, so

$$\langle v(x') - v(x), x' - x \rangle \geq 0 \tag{79}$$

implying first $\langle v(x') - v(x), x' - x \rangle = 0$ by monotonicity and then $\langle v(x'), x' - x \rangle = 0$. Thus, $x = x'$ by (U). $\square$

The regularized best response map is likewise defined as

$$\mathsf{BR}_{\varepsilon h}(x) = \arg\max_{x' \in \mathcal{X}} \{ \langle v(x), x' \rangle - \varepsilon h(x') \} \tag{80}$$

where $h \colon \mathcal{X} \to \mathbb{R}$ satisfies the same axioms as Definition 2.1. Note that $\mathsf{BR}_{\varepsilon h}$ satisfies (U) by construction.

Now, using these best-response maps, we may extend the definition of the dynamics (FP), (RFP) and (DA) to the more general setup considered here in the obvious way. We then have:

Proposition 6.2. *Suppose that* (U) *holds and let X_n, $n = 1, 2, \ldots$, be a sequence of states generated by* (FP). *If the accumulation points of $\bar{X}_n$ lie in $\mathrm{SVI}(\mathcal{X}, v)$, the same holds for the accumulation points of X_n.*

Proof. Let x^* be an accumulation point of $X_{n+1} = \mathsf{BR}(\bar{X}_n)$, and let X_{n_k+1} be a subsequence converging to x^*. By descending to a further subsequence if necessary, we may assume without loss of generality that $\bar{X}_{n_k}$ converges to some limit point $\bar{x}$. Furthermore, by the definition of (FP), we have:

$$\langle v(\bar{X}_{n_k}), X_{n_k+1} - x \rangle = \langle v(\bar{X}_{n_k}), \mathsf{BR}(\bar{X}_{n_k}) - x \rangle \geq 0 \quad \text{for all } x \in \mathcal{X}. \tag{81}$$

Hence, taking the limit $k \to \infty$, we get

$$\langle v(\bar{x}), x^* - x \rangle \geq 0 \quad \text{for all } x \in \mathcal{X}, \tag{82}$$

so $x^* \in \mathsf{BR}(\bar{x})$. Since $\bar{x} \in \mathrm{SVI}(\mathcal{X}, v)$ by assumption, we deduce further that $\bar{x} \in \mathsf{BR}(\bar{x})$. Hence, by (U), we conclude that $\bar{x} = x^*$, and our assertion follows. $\square$

Obviously, under (U), a similar property holds for regularized fictitious play.

6.5. **Preview for Part II.** Part II of our paper [17] deals with games with continuous action spaces, in the spirit of Example 6.1 (but with a continuous population of nonatomic players). The first step is to define the Stampacchia variational inequality when $\mathcal{A}$ is a compact metric space, Θ is an appropriate compact convex subset of $\Delta(\mathcal{A})$, v is a continuous map $\theta \mapsto v(\theta)[\cdot]$ from Θ to the space of continuous functions on $\mathcal{A}$, and the duality mapping is given by

$$\langle v(\theta), \nu \rangle = \int_{\mathcal{A}} v(\theta)[\alpha]\, d\nu[\alpha] \quad \text{for all } \theta, \nu \in \Theta. \tag{83}$$

In this framework, we introduce potential and monotone maps, and we prove the convergence of the discrete-time dynamics of fictitious play.

We show that this study covers the case of nonatomic games with (player-dependent) compact action spaces. Finally, we introduce a class of first-order mean-field games and prove that solutions of the Stampacchia variational inequality correspond to the equilibria of the game in normal form where $\mathcal{A}$ is the set of paths generated by the players.

Appendix A. **Properties of choice maps and the Fenchel coupling.** Our aim in this appendix is to provide some useful background and properties for the regularized choice map h and the Fenchel coupling.

To recall the basic setup, we assume below that h is a regularizer on $\mathcal{X}$ in the sense of Definition 2.1. The convex conjugate $h^* \colon \mathbb{R}^{\mathcal{A}} \to \mathbb{R}$ of h is then defined as

$$h^*(y) = \max_{x \in \mathcal{X}} \{\langle y, x \rangle - h(x)\}. \tag{A.1}$$

Since $\mathcal{X}$ is compact and h is strictly convex and continuous on $\mathcal{X}$, the maximum in (A.1) is attained at a unique point in $\mathcal{X}$, so $h^*(y)$ is finite for all $y \in \mathbb{R}^{\mathcal{A}}$ [1]. Moreover, by standard results in convex analysis [44, Chap. 26], h^* is differentiable on $\mathbb{R}^{\mathcal{A}}$ and its gradient satisfies the identity

$$\nabla h^*(y) = \arg\max_{x \in \mathcal{X}} \{\langle y, x \rangle - h(x)\} = Q(y) \tag{A.2}$$

where $Q \colon \mathbb{R}^{\mathcal{A}} \to \mathcal{X}$ is the choice map induced by h (cf. Definition 2.1).

The lemma below establishes the inverse differentiability property for h, namely that $\partial h = Q^{-1}$:

Lemma A.1. *Let h be a regularizer on $\mathcal{X}$. Then, for all $x \in \mathcal{X}$ and all $y \in \mathbb{R}^{\mathcal{A}}$, we have:*

$$x = Q(y) \iff y \in \partial h(x). \tag{A.3}$$

If, in addition, h satisfies (Diff), we also have

$$\langle \widetilde{\nabla} h(x), x - p \rangle \leq \langle y, x - p \rangle \tag{A.4}$$

for all $p \in \mathcal{X}$ and all $y \in \partial h(x)$, $x \in \operatorname{dom} \partial h$.

Proof of Lemma A.1. To prove (A.3), note that x solves (A.2) if and only if $y - \partial h(x) \ni 0$, i.e., if and only if $y \in \partial h(x)$.

For the inequality (A.4), it suffices to show it holds for all $p \in \operatorname{dom} \partial h$ (by continuity). To do so, let

$$\phi(t) = h(x + t(p - x)) - [h(x) + \langle y, x + t(p - x) \rangle]. \tag{A.5}$$

Since h is strictly convex and $y \in \partial h(x)$ by (A.3), it follows that $\phi(t) \geq 0$ with equality if and only if $t = 0$. Moreover, note that $\psi(t) = \langle \widetilde{\nabla} h(x + t(p - x)) - y, p - x \rangle$ is a continuous subgradient selection of ϕ. Given that ϕ and ψ are both continuous

on $[0, 1]$, it follows a fortiori that ϕ is continuously differentiable and $\phi' = \psi$ on $[0, 1]$. Thus, with ϕ convex and $\phi(t) \geq 0 = \phi(0)$ for all $t \in [0, 1]$, we conclude that $\phi'(0) = \langle \widetilde{\nabla} h(x) - y, p - x \rangle \geq 0$, from which our claim follows. $\square$

Lemma A.2. *Let h be a regularizer on $\mathcal{X}$. Then, for all $p \in \mathcal{X}$ and all $y \in \mathbb{R}^{\mathcal{A}}$, we have*

$$F_h(p, y) \geq 0 \quad \text{with equality if and only if } p = Q(y). \tag{A.6}$$

If, in addition, if h satisfies (StrCvx), we also have:

$$F_h(p, y) \geq \tfrac{K}{2} \|Q(y) - p\|^2. \tag{A.7}$$

Remark. In the same spirit as (A.7), if h is L-smooth, we also have the upper bound $(L/2)\|Q(y) - p\|^2 \geq F_h(p, y)$.

Proof. Our first claim is a trivial consequence of the Fenchel-Young inequality and the strict convexity of h. For our second claim, let $x = Q(y)$. Then, by the definition (A.1) of the convex conjugate of h, we have:

$$\begin{aligned}
F_h(p, y) &= h(p) + h^*(y) - \langle y, p \rangle \\
&= h(p) + \langle y, Q(y) \rangle - h(Q(y)) - \langle y, p \rangle \\
&= h(p) - h(x) - \langle y, p - x \rangle
\end{aligned} \tag{A.8}$$

Now, since h is K-strongly convex, we also have:

$$h((1 - \lambda)x + \lambda p) + \lambda(1 - \lambda)(K/2)\|x - p\|^2 \leq (1 - \lambda)h(x) + \lambda h(p) \tag{A.9}$$

and since $y \in \partial h(x)$ by (A.3), we get:

$$\langle y, x \rangle - h(x) \geq \langle y, (1 - \lambda)x + \lambda p \rangle - h((1 - \lambda)x + \lambda p). \tag{A.10}$$

Thus, after rearranging, dividing by λ, and letting $\lambda \to 0$, we obtain

$$h(p) - h(x) - \langle y, p - x \rangle \geq (K/2)\|x - p\|^2 \tag{A.11}$$

and our assertion follows from (A.8). $\square$

By virtue of this lemma, we obtain the following convergence criterion in terms of the Fenchel coupling:

Corollary A.1. *Let h be a regularizer on $\mathcal{X}$ satisfying (StrCvx), fix a base point $p \in \mathcal{X}$, and let y_n be a sequence in $\mathbb{R}^{\mathcal{A}}$. Then, if $F_h(p, y_n) \to 0$, we also have $Q(y_n) \to p$.*

Proof. By Lemma A.2, we have $\|Q(y_n) - p\| \leq \sqrt{(2/K)F_h(p, y_n)}$, so our claim follows. $\square$

REFERENCES

[1] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd edition, Springer, Cham, 2017.

[2] M. Benaïm, Dynamics of stochastic approximation algorithms, *Séminaire de Probabilités XXXIII*, **1709** (1999), 1–68.

[3] M. Benaïm and M. Faure, Consistency of vanishingly smooth fictitious play, *Math. Oper. Res.*, **38** (2013), 437–450.

[4] M. Benaïm, J. Hofbauer and S. Sorin, Stochastic approximations and differential inclusions, *SIAM J. Control Optim.*, **44** (2005), 328–348.

[5] M. Benaïm, J. Hofbauer and S. Sorin, Stochastic approximations and differential inclusions, part II: Applications, *Applications. Math. Oper. Res.*, **31** (2006), 673–695.

[6] C. Berge, *Topological Spaces*, Dover Publications, Inc., Mineola, NY, 1997.

[7] M. Bravo and P. Mertikopoulos, On the robustness of learning in games with stochastically perturbed payoff observations, *Games Econom. Behav.*, **103** (2017), 41–66.

[8] G. W. Brown, Iterative solutions of games by fictitious play, *Activity Analysis of Productions and Allocation*, (1951), 374–376.

[9] N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games*, Cambridge University Press, 2006.

[10] C. Conley, *Isolated Invariant Set and the Morse Index*, American Mathematical Society, Providence, RI, 1978.

[11] S. Dafermos, Traffic equilibrium and variational inequalities, *Transportation Sci.*, **14** (1980), 42–54.

[12] E. van Damme, *Stability and Perfection of Nash Equilibria*, Springer-Verlag, Berlin, 1987.

[13] D. Friedman, Evolutionary games in economics, *Econometrica*, **59** (1991), 637–666.

[14] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*, MIT Press, Cambridge, MA, 1998.

[15] D. Fudenberg and D. K. Levine, Conditional universal consistency, *Games and Economic Behavior*, **29** (1999), 104–130.

[16] I. Gilboa and A. Matsui, Social stability and equilibrium, *Econometrica*, **59** (1991), 859–867.

[17] S. Hadikhanloo, R. Laraki, P. Mertikopoulos and S. Sorin, *Learning in Nonatomic Games, Part II: Continuous action spaces and mean-field games.* mimeo, 2021.

[18] A. Héliou, M. Martin, P. Mertikopoulos and T. Rahier, Online non-convex optimization with imperfect feedback, In *NeurIPS '20: Proceedings of the 34th International Conference on Neural Information Processing Systems*, 2020.

[19] J. Hofbauer and W. H. Sandholm, On the global convergence of stochastic fictitious play, *Econometrica*, **70** (2002), 2265–2294.

[20] J. Hofbauer and W. H. Sandholm, Stable games and their dynamics, *J. Econom. Theory*, **144** (2009), 1665–1693.

[21] J. Hofbauer, P. Schuster and K. Sigmund, A note on evolutionarily stable strategies and game dynamics, *J. Theoret. Biol.*, **81** (1979), 609–612.

[22] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*, Cambridge University Press, Cambridge, UK, 1998.

[23] J. Hofbauer and K. Sigmund, Evolutionary game dynamics, *Bull. Amer. Math. Soc.*, **40** (2003), 479–519.

[24] J. Hofbauer, S. Sorin and Y. Viossat, Time average replicator and best reply dynamics, *Math. Oper. Res.*, **34** (2009), 263–269.

[25] A. Juditsky, A. S. Nemirovski and C. Tauvel, Solving variational inequalities with stochastic mirror-prox algorithm, *Stoch. Syst.*, **1** (2011), 17–58.

[26] J. Kwon, *Stratégies de Descente Miroir pour La Minimisation du Regret et L'approchabilité*, PhD thesis, Université Pierre-et-Marie-Curie, 2016.

[27] J. Kwon and P. Mertikopoulos, A continuous-time approach to online optimization, *J. Dyn. Games*, **4** (2017), 125–148.

[28] A. Mas-Colell, On a theorem of Schmeidler, *J. Math. Econom.*, **13** (1984), 201–206.

[29] J. Maynard Smith and G. R. Price, The logic of animal conflict, *Nature*, **246** (1973), 15–18.

[30] D. L. McFadden, Conditional logit analysis of qualitative choice behavior, *Frontiers in Econometrics*, (1974), 105–142.

[31] R. D. McKelvey and T. R. Palfrey, Quantal response equilibria for normal form games, *Games Econom. Behav.*, **10** (1995), 6–38.

[32] P. Mertikopoulos and A. L. Moustakas, The emergence of rational behavior in the presence of stochastic perturbations, *Ann. Appl. Probab.*, **20** (2010), 1359–1388.

[33] P. Mertikopoulos and W. H. Sandholm, Learning in games via reinforcement and regularization, *Math. Oper. Res.*, **41** (2016), 1297–1324.

[34] P. Mertikopoulos and W. H. Sandholm, Riemannian game dynamics, *J. Econom. Theory*, **177** (2018), 315–364.

[35] P. Mertikopoulos and M. Staudigl, On the convergence of gradient-like flows with noisy gradient input, *SIAM J. Optim.*, **28** (2018), 163–197.

[36] P. Mertikopoulos and Z. Zhou, Learning in games with continuous action sets and unknown payoff functions, *Math. Program.*, **173** (2019), 465–507.

[37] G. J. Minty, On the generalization of a direct method of the calculus of variations, *Bull. Amer. Math. Soc.*, **73** (1967), 315–321.

[38] D. Monderer and L. S. Shapley, Potential games, *Games Econom. Behav.*, **14** (1996), 124–143.

[39] D. Monderer and L. S. Shapley, Fictitious play property for games with identical interests, *J. Economic Theory*, **68** (1996), 258–265.

[40] A. Nemirovski, Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems, *SIAM J. Optim.*, **15** (2004), 229–251.

[41] Y. Nesterov, Dual extrapolation and its applications to solving variational inequalities and related problems, *Math. Program.*, **109** (2007), 319–344.

[42] Y. Nesterov, Primal-dual subgradient methods for convex problems, *Math. Program.*, **120** (2009), 221–259.

[43] J. Robinson, An iterative method for solving a game, *Ann. of Math.*, **54** (1951), 296–301.

[44] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, N. J., 1970.

[45] A. Rustichini, Optimal properties of stimulus-response learning models, *Games Econom. Behav.*, **29** (1999), 244–273.

[46] W. H. Sandholm, Potential games with continuous player sets, *J. Econom. Theory*, **97** (2001), 81–108.

[47] W. H. Sandholm, *Population Games and Evolutionary Dynamics*, MIT Press, Cambridge, MA, 2010.

[48] W. H. Sandholm, Population games and deterministic evolutionary dynamics, *Handbook of Game Theory IV*, (2015), 703–778.

[49] D. Schmeidler, Equilibrium points of nonatomic games, *J. Statist. Phys.*, **7** (1973), 295–300.

[50] S. Shalev-Shwartz, Online learning and online convex optimization, *Foundations and Trends in Machine Learning*, **4** (2011), 107–194.

[51] S. Shalev-Shwartz and Y. Singer, Convex repeated games and Fenchel duality, In *NIPS' 06: Proceedings of the 19th Annual Conference on Neural Information Processing Systems*, (2006), 1265–1272.

[52] S. Sorin, Exponential weight algorithm in continuous time, *Math. Program.*, **116** (2009), 513–528.

[53] S. Sorin, Replicator dynamics: Old and new, *J. Dyn. Games*, **7** (2020), 365–386.

[54] S. Sorin, No-regret algorithms in online learning, games and convex optimization, Mimeo, 2021.

[55] S. Sorin and C. Wan, Finite composite games: Equilibria and dynamics, *J. Dyn. Games*, **3** (2016), 101–120.

[56] P. D. Taylor, Evolutionarily stable strategies with two types of player, *J. Appl. Probab.*, **16** (1979), 76–83.

[57] P. D. Taylor and L. B. Jonker, Evolutionary stable strategies and game dynamics, *Math. Biosci.*, **40** (1978), 145–156.

[58] Jörgen W. Weibull, *Evolutionary Game Theory*, MIT Press, Cambridge, MA, 1995.

E-mail address: saeed.hadikhanloo@tweag.io

E-mail address: rida.laraki@dauphine.fr

E-mail address: panayotis.mertikopoulos@imag.fr

E-mail address: sylvain.sorin@imj-prg.fr