Invariance and Concentration Properties of Gradient-based Learning in Games

Kyriakos Lotidis*, Panayotis Mertikopoulos♯, Nicholas Bambos§ and Jose Blanchet†

Abstract— **In this paper, we study the long-run behavior of learning in strongly monotone games with stochastic, gradient-based feedback. For concreteness, we focus on the *stochastic projected gradient* (SPG) algorithm, and we examine the asymptotic distribution of its iterates when the method is run with constant step-size updates (the de facto choice for practical deployments of the algorithm). In contrast to variants of the method with a vanishing step-size case, SPG with a constant step-size does not converge: instead, it reaches a neighborhood of the game's Nash equilibrium at an exponential rate, and then, due to persistent noise, it fluctuates in its vicinity without converging (occasionally moving away on rare occasions). We provide a theoretical quantification of this behavior by analyzing the Markovian structure of the process. Namely, we show that, regardless of the algorithm's initialization, the distribution of its iterates converges at a geometric rate to a unique invariant measure which is concentrated in a neighborhood of the game's Nash equilibrium. More explicitly, we quantify the degree of this concentration and the rate of convergence of the algorithm's empirical frequency of play to the invariant measure of the process in Wasserstein distance, and we provide explicit bounds in terms of the method's step-size, the variance of the noise entering the process, and the geometric features of the game's payoff landscape.**

I. Introduction

Owing to its simplicity and empirical successes, *stochastic gradient descent* (SGD) has become the de facto method for solving large-scale optimization problems and training a wide range of machine learning (ML) models and architectures. This is especially the case in multi-agent models and applications – from adversarial machine learning [1] and generative models [2], [3], to multi-agent reinforcement learning [4] and power control [5], [6] – where the co-existence and interference of multiple interacting agents precludes the use of more sophisticated methods (like Newton or LBFGS-type updates).

The standard paradigm for learning in multi-agent environments is the framework of online learning in games [7]–[9] which roughly unfolds as follows:

1) At each stage of a repeated multi-agent decision process, every participating player selects an action.
2) Each player receives a reward determined by the action they chose and the actions chosen by all other players.
3) Based on their payoffs and any other feedback, the players update their actions, and the process repeats.

In this general context, the multi-agent incarnation of SGD boils down to *stochastic projected gradient* (SPG) updates of the form

$$X_{n+1} = \Pi(X_n + \gamma_n V_n) \tag{SPG}$$

where, deferring a detailed presentation for later, $n = 1, 2, \dots$ denotes the running index of the process, X_n is the players' action profile at time n, V_n is the players' stochastic gradient feedback, γ_n is the algorithm's step-size parameter, and Π is a projection operator.

In a multi-agent setting as above, the convergence and stability analysis of the algorithm becomes significantly more intricate than the vanilla, single-agent case. Indeed, unlike single-agent convex optimization problems, multi-agent learning dynamics may exhibit complex behaviors such as cycles, divergence, or recurrence [10]–[12] – and this, even when the game is *monotone* (the multi-agent analogue of convexity). For this reason, a standard case study for analyzing the properties of (SPG) from a theoretical standpoint is that of *strongly monotone* games (the multi-agent analogue of *strong* convexity), which have a range of desirable features: *a*) when the players' action sets are compact, strongly monotone games admit a *unique* Nash equilibrium [13]; *b*) this Nash equilibrium can be characterized as the unique solution of a suitable (strongly monotone) variational inequality [13]; and *c*) under certain hypotheses, it is possible to show that gradient-based methods lead to a contraction principle, which can then be exploited to yield convergence in a wide range of learning settings.

In the case of (SPG) applied to strongly monotone games – or, more generally, to strongly monotone variational inequalities (VIs) – the sequence of play X_n is known to converge almost surely [9], [14]–[16] under a diminishing step-size sequence γ_n, $n = 1, 2, \dots$, satisfying the so-called Robbins–Monro (or "$L^2 - L^1$") summability conditions

$$\sum_{n=0}^{\infty} \gamma_n = \infty, \quad \sum_{n=0}^{\infty} \gamma_n^2 < \infty \tag{1}$$

In practical applications and deployments of (SPG), it is common to adopt a *constant* step-size implementation for several reasons. First, constant step-sizes are simpler to deploy and maintain, since the tuning process is easier and more

*Dept. of Management Science & Engineering, Stanford University, klotidis@stanford.edu.

♯Univ. Grenoble-Alpes, CNRS, Inria, Grenoble INP, LIG, 38000, France panayotis.mertikopoulos@imag.fr.

§Dept. of Electrical Engineering and Management Science & Engineering, Stanford University, bambos@stanford.edu.

†Dept. of Management Science & Engineering, Stanford University, jose.blanchet@stanford.edu.

J. Blanchet gratefully acknowledges support from DoD through the grants Air Force Office of Scientific Research under award number FA9550-20-1-0397 and ONR 1398311, also support from NSF via grants 2229012, 2312204, 2403007 is gratefully acknowledged. P. Mertikopoulos is also a member of the Archimedes Research Unit/Athena RC and acknowledges support by the French National Research Agency (ANR) in the framework of the PEPR IA FOUNDRY project (ANR-23-PEIA-0003), the "Investissements d'avenir" program (ANR-15-IDEX-02), and the LabEx PERSYVAL (ANR-11-LABX-0025-01), under grant IRGA-SPICE (G7H-IRG24E90), and project MIS 5154714 of the National Recovery and Resilience Plan Greece 2.0 funded by the European Union under the NextGenerationEU Program.

robust. From a practical standpoint, gradient methods with a vanishing step-size typically suffer from long warm-up periods and slow convergence to a neighborhood of the equilibrium point. By contrast, constant step-size methods in machine learning settings reach the vicinity of a solution much faster, often within 0.1% accuracy [34]. Many state-of-the-art architectures, including transformers and large language models, employ step-size schedules that remain effectively constant over billions, or even trillions of samples [**?**].

At the same time, the use of a constant step-size does not come without its disadvantages: Unlike their diminishing-step-size counterparts, constant step-size implementations *do not* converge to a Nash equilibrium (or, more generally, the solution of the associated VI), but instead end up fluctuating in its vicinity due to the effects of persistent stochastic noise.[1] In such settings, without more stringent assumptions in place (like interpolation or vanishing noise), it is more appropriate to replace questions of convergence with questions of *concentration* –namely, asking where the iterates tend to spend most of their time and how sharply they concentrate around particular the game's equilibrium regions.

This raises the natural question:

What is the long-run behavior of stochastic gradient-based learning with a constant step-size, and what structure, if any, governs its asymptotic properties?

To answer this, a central object of our paper is the *invariant measure* of the Markov process induced by (SPG). Specifically, the invariant measure captures the long-run statistical behavior of the algorithm, as it describes the distribution toward which the iterates stabilize in a probabilistic sense, even when they fail to converge in a pathwise manner. Thus, rather than focusing on pointwise convergence, we study the invariant measure as a principled tool for characterizing the stationary behavior of constant step-size gradient-based learning methods in strongly monotone games, allowing us to quantify how the iterates concentrate around the attractor of the dynamics.

Our contributions and related work: A broad line of work studies the computation of Nash equilibria in monotone games through the lens of VI. Classical results provide convergence guarantees for monotone and strongly monotone problems under various projection and extragradient methods [17]–[19]. In the stochastic regime, convergence in expectation, high-probability bounds and last-iterate almost sure convergence are established under diminishing step-size schedules [9], [14], [15], [20]–[30]. More relevant to our work is the recent paper of [31] on weak quasi strongly monotone VIs, which studies the ergodic properties of the constant step-size variants of the *stochastic extra-gradient* and *stochastic gradient desscent-ascent* algorithms, under the assumption that the noise possesses a density component that is uniformly positive. This condition enables the use of classical ergodic results for Harris recurrent and positive recurrent Markov chains, ensuring irreducibility and geometric convergence in total variation [32, Theorem 15.0.1]. In contrast, we do not impose any density assumptions on the noise to establish the uniqueness of the invariant measure via ergodicity or recurrence properties. Instead, we leverage the strong monotonicity of the problem to construct a stochastic contraction argument. This approach allows us to establish the uniqueness of the invariant measure and further yields geometric convergence of the iterates to this measure in Wasserstein distance.

Furthermore, [33] study the ergodic properties of constant step-size SGD in the minimization of a non-smooth non-convex objective function, under similar density assumptions to [31], while [**?**], [**?**] analyze the long-run and global convergence time of SGD in non-convex problems. In the convex regime, [34] studies constant step-size SGD for strongly convex and smooth functions, showing convergence to the unique stationary distribution and providing an explicit asymptotic expansion of the moments of the averaged SGD iterates. Finally, several results in the optimization literature have established the asymptotic normality of the *averaged* iterates of SGD [35]–[37]. However, in the game-theoretic regime, we do not adopt a computational viewpoint; instead, the primary focus is on the convergence of the *actual sequence of play*, rather than on its weighted averages. To our knowledge, the only study of the invariant distribution of gradient-based learning in games has focused on continuous-time models, cf. [**?**], [**?**], [**?**], [**?**] and references therein.

Paper outline: Our paper is organized as follows. In Section II we introduce some notions of game-theoretic learning and Markov processes required for our results, while in Section III we present our learning dynamics. In Section IV, present our main results regarding the uniqueness of an invariant measure and the convergence to it. Finally, in Section V we provide numerical simulations to illustrate the behavior of the invariant measure with respect to changes in the step-size and the level of uncertainty.

II. Preliminaries

We start by briefly reviewing some basics of continuous games and Markov processes, introducing the necessary context for our results.

Continuous games: A *continuous game* consists of the following primitives:

1) A finite set of *players* $i \in \mathcal{N} = \{1, \dots, N\}$.
2) Each player $i \in \mathcal{N}$ has access to a compact convex set $\mathcal{X}_i$ of a finite dimensional vector space $\mathcal{V}_i$, describing the set of actions available to the player. We will write $\mathcal{X} \coloneqq \prod_i \mathcal{X}_i$ for the space of all ensembles $x = (x_1, \dots, x_N)$ of actions $x_i \in \mathcal{X}_i$ that are independently chosen by each player $i \in \mathcal{N}$. We will write $x = (x_i; x_{-i})$ to emphasize the action of player $i \in \mathcal{N}$ against the action collection $x_{-i} = (x_j)_{j \neq i}$ of the rest.
3) The players' rewards are determined by their individual *payoff functions* $u_i \colon \mathcal{X} \to \mathbb{R}$, which are assumed to be continuously differentiable in x_i. Moreover, we will denote the individual gradient vector of player $i \in \mathcal{N}$ by

$$v_i(x) \equiv \nabla_{x_i} u_i(x_i; x_{-i}) \tag{2}$$

[1] In the absence of noise however, convergence is assured as long as the method's step-size is less than the game's inverse smoothness modulus.

and the collection $v(x) \equiv (v_1(x), \dots, v_N(x))$ thereof.

A *continuous game* is then defined as a tuple $\mathcal{G} \equiv \mathcal{G}(\mathcal{N}, \mathcal{X}, u)$ with players $\mathcal{N}$, actions $\mathcal{X}$, and payoff functions $u \equiv (u_i)_{i\in\mathcal{N}}$ as described above.

***Nash equilibrium**:* The classical solution concept of a *Nash equilibrium* (NE) characterizes the actions $x^* \in \mathcal{X}$ from which no player has incentive to unilaterally deviate. More rigorously, $x^* \in \mathcal{X}$ is a NE if

$$u_i(x^*) \geq u_i(x_i; x^*_{-i}) \quad \text{for all } x_i \in \mathcal{X}_i, i \in \mathcal{N}. \tag{NE}$$

***Monotone games**:* A continuous game $\mathcal{G}$ is called *monotone* if the following monotonicity condition holds

$$\langle v(x) - v(x'), x - x' \rangle \leq 0 \quad \text{for all } x, x' \in \mathcal{X}. \tag{3}$$

Setting $x_{-i} = x'_{-i}$ for each $i \in \mathcal{N}$, we readily obtain that the individual payoff function u_i of each player $i \in \mathcal{N}$ is concave in its actions, i.e., the function $x_i \mapsto u_i(x_i; x_{-i})$. Standard arguments of convex analysis [38] show that the Nash equilibria of a monotone game $\mathcal{G}$ are precisely the solutions of the variational inequality

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

so, the existence of a NE point follows from standard results. We will use this equivalence freely in the sequel. Finally, a game $\mathcal{G}$ is called *λ-strongly monotone* if there exists $\lambda > 0$ such that

$$\langle v(x) - v(x'), x - x' \rangle \leq -\lambda \|x - x'\|_2^2 \quad \text{for all } x, x' \in \mathcal{X}. \tag{4}$$

It is well known [18] that a strongly monotone game $\mathcal{G}$ has a unique NE point $x^* \in \mathcal{X}$.

***Wasserstein distance & Markov processes**:* We close this section by introducing the *Wasserstein-p* distance and some notions on *Markov processes*. In what follows, we denote by $\mathcal{P}(\mathcal{X})$ the set of probability measures on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, where $\mathcal{B}(\mathcal{X})$ is the Borel σ-algebra on $\mathcal{X}$.

Specifically, a *Markov kernel* $P : \mathcal{X} \times \mathcal{B}(\mathcal{X}) \to [0, 1]$ is a function such that (i) for any $A \in \mathcal{B}(\mathcal{X})$ the map $x \mapsto P(x, A)$ is $\mathcal{B}(\mathcal{X})$-measurable, and (ii) for any $x \in \mathcal{X}$, the map $A \mapsto P(x, A)$ is a probability measure on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. A random process $(X_n)_{n\in\mathbb{N}}$ on $\mathcal{X}$ is a time-homogeneous Markov chain with kernel P if for all $A \in \mathcal{B}(\mathcal{X})$, $\mathbb{P}(X_{n+1} \in A \mid X_0, \dots, X_n) = \mathbb{P}(X_{n+1} \in A \mid X_n) = P(X_n, A)$. Moreover, for a random process $(X_n)_{n\in\mathbb{N}}$, we will denote the law of X_k given $X_0 = x$ as $P_x^{(k)}$, i.e.,

$$P_x^{(k)}(A) = \mathbb{P}(X_k \in A \mid X_0 = x) \quad \text{for any } A \in \mathcal{B}(\mathcal{X}). \tag{5}$$

To introduce the concept of the invariant measure, we begin by defining the action of a kernel on a probability measure. Given $\mu \in \mathcal{P}(\mathcal{X})$ and a Markov kernel P, the probability measure $\mu P \in \mathcal{P}(\mathcal{X})$ is defined as

$$\mu P(A) = \int \mu(dx) P(x, A) \quad \text{for any } A \in \mathcal{B}(\mathcal{X}) \tag{6}$$

and captures the probability that the next state lies in A if the current is sampled from μ. With this definition in hand, $\mu \in \mathcal{P}(\mathcal{X})$ will be called an *invariant measure* on $\mathcal{X}$ if $\mu = \mu P$. Note, that if μ is an invariant measure for a Markov kernel P, and $X_0 \sim \mu$, then $X_n \sim \mu$ for all n, as well. Intuitively, a probability measure over $\mathcal{X}$ is invariant if it stays the same under the underlying dynamics.

Furthermore, for two probability measures $\mu, \nu \in \mathcal{P}(\mathcal{X})$ and $p \geq 1$, the *Wasserstein-p* distance [39] is defined as

$$\mathcal{W}_p(\mu, \nu) = \inf_{\pi \in \Gamma(\mu, \nu)} \left(\int \|x - y\|_2^p \, d\pi(x, y) \right)^{1/p} \tag{7}$$

where $\Gamma(\mu, \nu)$ denotes the set of probability measures (known also as couplings) on $\mathcal{X} \times \mathcal{X}$ with marginals $\mu, \nu \in \mathcal{P}(\mathcal{X})$. Additionally, for compact spaces, the *Wasserstein-p* distance metrizes the weak convergence of probability measures [39], meaning that $\mathcal{W}_p(\mu_n, \mu) \to 0$ if and only if $\mu_n \Rightarrow \mu$, where '$\Rightarrow$' stands for the weak convergence of probability measures.

We are now ready to present the main learning dynamics studied in this paper.

III. Learning Dynamics

One of the most widely used algorithm in continuous games is the so-called *stochastic projected gradient* (SPG) algorithm which, in our notation, unfolds as

$$X_{i,n+1} = \Pi_i\left(X_{i,n} + \gamma V_{i,n}\right) \quad \text{for all } i \in \mathcal{N}, n \in \mathbb{N} \tag{SPG}$$

where $\gamma > 0$ is the step-size hyperparameter. The operator $\Pi_i : \mathcal{X}_i \to \mathcal{X}_i$ is the Euclidean projection operator, defined as $\Pi_i(x) = \arg\min_{x' \in \mathcal{X}_i} \|x - x'\|_2^2$, and which is a non-expansive operator [40], i.e.,

$$\|\Pi_i(x) - \Pi_i(x')\|_2 \leq \|x - x'\|_2 \tag{8}$$

We further make the following two blanket assumptions, which we adopt for the remainder of the paper.

***Assumption 1**:* We consider a λ-strongly monotone game $\mathcal{G}(\mathcal{N}, \mathcal{X}, u)$, which implies the uniqueness of a NE $x^* \in \mathcal{X}$.

***Assumption 2**:* At every time step n, each player $i \in \mathcal{N}$ has access to a *stochastic first-order oracle* (SFO), observing a noisy version of their individual gradient vector

$$V_{i,n} \equiv V_i(X_n, \omega_n) \tag{9}$$

where $\{\omega_n\}_{n\in\mathbb{N}}$ are sampled independently from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and independent from the initial state X_0. Moreover, the functions $x \mapsto V_i(x, \omega)$ are assumed to be *continuous*, with the following statistical properties:

(i) Conditionally unbiased: $\mathbb{E}[V_i(X_n, \omega) \mid \mathcal{F}_n] = v_i(X_n)$

(ii) Bounded in mean square: $\mathbb{E}[\|V_i(X_n, \omega)\|_2^2 \mid \mathcal{F}_n] \leq \sigma^2$

(iii) Lipschitz in the mean:

$$\mathbb{E}\left[\|V_i(x, \omega) - V_i(x', \omega)\|_2^2\right] \leq L\|x - x'\|_2^2 \; \forall x, x' \in \mathcal{X}$$

where $(\mathcal{F}_n)_{n\in\mathbb{N}}$ is the natural filtration associated to $(X_n)_{n\in\mathbb{N}}$, i.e., $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$.

Finally, denoting the ensemble of projections as $\Pi(x) \equiv (\Pi_1(x_1), \ldots, \Pi_N(x_N))$, we will write (SPG) as $X_{n+1} = \Pi(X_n + \gamma V_n)$ in the sequel.

Markovian structure of* (SPG)*: To conclude this section, we note that the sequence of play $(X_n)_{n\in\mathbb{N}}$ generated by the (SPG) dynamics is a Markov chain on $\mathcal{X}$. For this, let the function $F : \mathcal{X} \times \Omega \to \mathcal{X}$ be defined as

$$F(x, \omega) \equiv \Pi(x + \gamma V(x, \omega)) \tag{10}$$

Then, the (SPG) dynamics can be described by the time-homogeneous equation of the form

$$X_{n+1} = F(X_n, \omega_n) = \Pi(X_n + \gamma V(X_n, \omega_n)) \tag{11}$$

with ω_n sampled from $\mathbb{P}$ independent for each n, verifying that $(X_n)_{n\in\mathbb{N}}$ is a time-homogeneous Markov chain [41].

In the sequel, we exploit the Markovian structure of the iterates $(X_n)_{n\in\mathbb{N}}$ from (SPG) to derive our main results on the existence, uniqueness, and convergence properties of the invariant measure.

IV. Analysis and Main Results

In this section, we establish the existence and uniqueness of the invariant measure, as well as the convergence to it.

A. Existence and uniqueness of the invariant measure

In order to meaningfully discuss the long-run behavior of the dynamics, it is essential to show the existence of an invariant measure, as it forms the foundation for describing their stationary properties. Specifically, we have:

Theorem 1. *The Markov process $(X_n)_{n\in\mathbb{N}}$ generated by the* (SPG) *admits an invariant probability measure μ in $\mathcal{X}$.*

The proof of Theorem 1 relies on the, rather technical, (weak) Feller property of Markov processes. Specifically, a Markov process $(X_n)_{n\in\mathbb{N}}$ on $\mathcal{X}$ with kernel P is (weak) Feller if for any bounded and continuous function $g : \mathcal{X} \to \mathbb{R}$, the function $x \mapsto \int P(x, dy)g(y)$ is bounded and continuous [42]. It is a topological condition that essentially ensures a form of regularity in how the process evolves. In simple words, if the process is initialized from nearby points, the expected future value under bounded continuous functions remain close.

In the next proposition, we show that the (SPG) iterates satisfy the (weak) Feller property.

Proposition 1. *The random process $(X_n)_{n\in\mathbb{N}}$ generated by the dynamics* (SPG) *is a (weak) Feller Markov chain in $\mathcal{X}$.*

Proof. First, note that the function $x \mapsto F(x, \omega)$ defined in (10) is continuous in $\mathcal{X}$ for all $\omega \in \Omega$, since for any $x \in \mathcal{X}$ and any sequence $\{x_k\}_{k\in\mathbb{N}} \subseteq \mathcal{X}$ with $x_k \to x$, it holds:

$$\lim_{k\to\infty} F(x_k, \omega) = F(x, \omega) \tag{12}$$

invoking the non-expansiveness property of the projection operator (8), and the continuity of $x \mapsto V(x, \omega)$.

Now, let $g : \mathcal{X} \to \mathbb{R}$ be a continuous and bounded function. Our goal is to show that the function

$$x \mapsto \mathbb{E}_x[g(X_1)] \tag{13}$$

is continuous and bounded, where $X_1 = F(x, \omega_1)$. For this, let $x \in \mathcal{X}$ and let any sequence $\{x_k\}_{k\in\mathbb{N}} \subseteq \mathcal{X}$ with $x_k \to x$. Since $F(\cdot, \omega_1)$ is continuous for all $\omega_1 \in \Omega$, as we showed previously, and g is assumed to be continuous, we obtain:

$$\lim_{k\to\infty} g(F(x_k, \omega_1)) = g(F(x, \omega_1)) \quad \mathbb{P}\text{-a.s.} \tag{14}$$

Finally, since g is bounded, by the dominated convergence theorem, we obtain:

$$\lim_{k\to\infty} \mathbb{E}_{x_k}[g(X_1)] = \lim_{k\to\infty} \mathbb{E}[g(F(x_k, \omega_1))] \tag{15}$$

$$= \lim_{k\to\infty} \int g(F(x_k, \omega_1))\, \mathbb{P}(d\omega_1) \tag{16}$$

$$= \int g(F(x, \omega_1))\, \mathbb{P}(d\omega_1) \tag{17}$$

$$= \mathbb{E}_x[g(X_1)] \tag{18}$$

which proves the continuity of $x \mapsto \mathbb{E}_x[g(X_1)]$. The boundedness of the function follows by the boundedness of f, and our proof is complete. ■

Having established the (weak) Feller property of the (SPG) iterates, we are ready to proceed to the proof of Theorem 1.

Proof of Theorem 1. Since $(X_n)_{n\in\mathbb{N}}$ is a (weak) Feller Markov chain and $\mathcal{X}$ is compact, the result follows from the Krylov–Bogolyubov theorem [44], which states that if the sequence of the probability laws $(P_x^{(n)})_{n\in\mathbb{N}}$ is tight for some initial state $x \in \mathcal{X}$, then $(X_n)_{n\in\mathbb{N}}$ admits at least one invariant measure. Thus, since any family of probability measures on a compact space is tight, and $\mathcal{X}$ is compact, the result follows. ■

The next result shows that the sequence of play $(X_n)_{n\in\mathbb{N}}$ admits a *unique invariant measure* for appropriate choices of the step-size, enabling a well-defined description of the system's stationary behavior.

Theorem 2. *The Markov process $(X_n)_{n\in\mathbb{N}}$ generated by the* (SPG) *dynamics for $\gamma < 2\lambda/NL$ admits a unique invariant probability measure $\mu \in \mathcal{P}(\mathcal{X})$.*

Proof. The idea of this proof relies on a stochastic contraction argument. More specifically, according to [45, Theorem 4.25], a Markov process defined via the time-homogeneous dynamics equation (11) can have at most one invariant probability measure if there exists $\alpha \in (0, 1)$ with

$$\mathbb{E}[\|F(x, \omega) - F(y, \omega)\|_2] \leq \alpha\|x - y\|_2 \quad \text{for all } x, y \in \mathcal{X}. \tag{19}$$

Specifically, in our case, for $x, y \in \mathcal{X}$, and using (8), we have:

$$\begin{aligned}\|F(x, \omega) - F(y, \omega)\|_2^2 &\leq \|(x + \gamma V(x, \omega)) - (y + \gamma V(y, \omega))\|_2^2 \\ &= \|x - y\|_2^2 + \gamma^2\|V(x, \omega) - V(y, \omega)\|_2^2\end{aligned}$$

$$+ 2\gamma\langle V(x,\omega) - V(y,\omega), x - y\rangle \tag{20}$$

Taking expectation with respec to ω, and using Assumption 2(iii) and the λ-strong monotonicity, we get

$$\begin{aligned}\mathbb{E}\big[\|F(x,\omega) - F(y,\omega)\|_2^2\big] &\le (1-2\gamma\lambda)\|x-y\|_2^2 \\ &\quad + \gamma^2\,\mathbb{E}\big[\|V(x,\omega) - V(y,\omega)\|_2^2\big] \\ &\le (1-2\gamma\lambda+\gamma^2 NL)\|x-y\|_2^2\end{aligned} \tag{21}$$

Note that the Lipschitz in the mean condition [Assumption 2(iii)] along with Jensen's inequality and the λ-strong monotonicity, guarantee that $\lambda^2 < NL$, and, thus, $(1-2\gamma\lambda+\gamma^2 NL) > 0$. Using Jensen's inequality, we obtain:

$$\begin{aligned}\mathbb{E}[\|F(x,\omega) - F(y,\omega)\|_2] &\le \sqrt{\mathbb{E}\big[\|F(x,\omega) - F(y,\omega)\|_2^2\big]} \\ &\le \sqrt{1-2\gamma\lambda+\gamma^2 NL}\,\|x-y\|_2\end{aligned} \tag{22}$$

Finally, for $\gamma < 2\lambda/NL$, we have $\sqrt{1-2\gamma\lambda+\gamma^2 NL} \in (0,1)$ which proves (19). ■

B. Convergence to the invariant measure

After establishing the existence and uniqueness of the invariant measure, we turn to studying the system's long-term behavior and convergence. In view of this, a central object in the study of ergodic systems is the mean occupation measure, $A \mapsto \mu_n(x, A)$, defined as

$$\mu_n(x,A) \equiv \frac{1}{n}\,\mathbb{E}_x\left[\sum_{k=0}^{n-1}\mathbb{1}\{X_k \in A\}\right] = \frac{1}{n}\sum_{k=0}^{n-1} P_x^{(k)}(A) \tag{23}$$

for any $x \in \mathcal{X}$, which plays a key role in quantifying the long-run behavior of the process. In particular, ergodic theorems guarantee that time averages, such as those captured by μ_n, converge to space averages under the invariant measure, forgetting the initial state $x \in \mathcal{X}$ and revealing how empirical behavior reflects the system's stationary properties. More concretely, we have:

Theorem 3. *Let the Markov process $(X_n)_{n\in\mathbb{N}}$ generated by the (SPG) for $\gamma < 2\lambda/NL$. Then, the mean occupation measure converges strongly, to the invariant measure μ for μ-a.e $x \in \mathcal{X}$, i.e., for all $A \in \mathcal{B}(\mathcal{X})$*

$$\lim_{n\to\infty} \mu_n(x,A) = \mu(A) \quad \textit{for } \mu\textit{-a.e } x \in \mathcal{X} \tag{24}$$

Moreover, for any $\delta > 0$ and $\gamma < \min\{2\lambda/NL, 2\lambda\delta^2/N\sigma^2\}$ the probability mass of μ around the ball $\mathbb{B}_\delta \equiv \{x \in \mathcal{X} : \|x - x^*\|_2^2 \le \delta\}$, is bounded by:

$$\mu(\mathbb{B}_\delta) \ge 1 - \frac{\gamma N\sigma^2}{2\lambda\delta^2} \tag{25}$$

Proof. Let some arbitrary $A \in \mathcal{B}(\mathcal{X})$. For the first part of the theorem, we will invoke the Birkhoff's individual ergodic theorem [41, Theorem 2.3.4], according to which the expectation of the indicator function $\mathbb{1}_A$ under the probability measure $\mu_n(x,\cdot)$ converges for μ-a.e. initial condition $x \in \mathcal{X}$. Under the special case of a unique invariant measure, the limit is constant for μ-a.e. initial condition $x \in \mathcal{X}$ and equal to the expectation of the function (i.e., $\mathbb{1}_A$) with respect to the invariant measure, $\int \mathbb{1}_A\, d\mu = \mu(A)$ [41, Proposition 2.4.2].

Thus, we readily get that

$$\lim_{n\to\infty} \mu_n(x,A) = \mu(A) \quad \mu\text{-a.e.} \tag{26}$$

for all $A \in \mathcal{B}(\mathcal{X})$, and the strong convergence of the mean occupation measure is established.

Now, regarding the second part of the theorem, for the running distance $\|X_n - x^*\|_2^2$, we obtain

$$\begin{aligned}\|X_{n+1} - x^*\|_2^2 &= \|\Pi(X_n + \gamma V_n) - \Pi(x^*)\|_2^2 \\ &\le \|X_n + \gamma V_n - x^*\|_2^2 \\ &\le \|X_n - x^*\|_2^2 + \gamma^2\|V_n\|_2^2 + 2\gamma\langle V_n, X_n - x^*\rangle\end{aligned}$$

Conditioning on X_n, the above inequality becomes:

$$\begin{aligned}\mathbb{E}\big[\|X_{n+1} - x^*\|_2^2 \,\big|\, X_n\big] &\le \|X_n - x^*\|_2^2 + \gamma^2\,\mathbb{E}\big[\|V_n\|_2^2 \,\big|\, X_n\big] \\ &\quad + 2\gamma\langle v(X_n), X_n - x^*\rangle \\ &\le \|X_n - x^*\|_2^2 + \gamma^2 N\sigma^2 \\ &\quad - 2\gamma\lambda\|X_n - x^*\|_2^2\end{aligned} \tag{27}$$

Thus, taking expectation conditional on $X_0 = x$ in both sides and iterating over n time steps, we obtain

$$\begin{aligned}\mathbb{E}_x\big[\|X_n - x^*\|_2^2\big] &\le \|x - x^*\|_2^2 + \gamma^2 N\sigma^2 n \\ &\quad - 2\gamma\lambda\,\mathbb{E}_x\left[\sum_{k=0}^{n-1}\|X_k - x^*\|_2^2\right]\end{aligned} \tag{28}$$

Now, for μ-a.e. $x \in \mathcal{X}$, we have

$$\begin{aligned}\mu_n(x, \mathbb{B}_\delta^c) &= \frac{\mathbb{E}_x\left[\sum_{k=0}^{n-1}\mathbb{1}\{X_k \notin \mathbb{B}_\delta\}\right]}{n} \\ &\le \frac{\mathbb{E}_x\left[\sum_{k=0}^{n-1}\|X_k - x^*\|_2^2\right]}{n\delta^2} \\ &\le \frac{1}{2\gamma\lambda n\delta^2}\|x^* - x\|_2^2 + \frac{\gamma N\sigma^2}{2\lambda\delta^2}\end{aligned} \tag{29}$$

Thus, taking $n \to \infty$, and using the strong convergence of the mean occupation measure, established before, we readily obtain that $\mu(\mathbb{B}_\delta^c) \le \frac{\gamma N\sigma^2}{2\lambda\delta^2}$ and our claim follows. ■

One important remark is that the strong convergence of the mean occupation measure (23) holds for μ-a.e. $x \in \mathcal{X}$. Without further assumptions on the structure of the noise, it is difficult to infer the support of the invariant measure μ. For instance, it can be shown [46, Lemma 2.2] that if the noise has density with full support, every state $x \in \mathcal{X}$ is accessible, and, subsequently, the support of μ consists of the whole space $\mathcal{X}$. However, we can show that both the law of the Markov process $\{X_n\}_{n\in\mathbb{N}}$, and the mean occupation measure (23), converge in Wasserstein distance for any initial state $x \in \mathcal{X}$. Formally, we have the following results.

Theorem 4. *Let the Markov process $(X_n)_{n\in\mathbb{N}}$ generated by (SPG) with $\gamma < 2\lambda/NL$. Then, the law $P_x^{(n)}$ of the process* *given $X_0 = x$, converges geometrically to the invariant measure μ in Wasserstein-1 distance, i.e.,*

$$\mathcal{W}_1(P_x^{(n)}, \mu) = \mathcal{O}\left(\rho^{n/2}\right) \tag{30}$$

for $\rho \equiv \left(1 - 2\gamma\lambda + \gamma^2 NL\right) \in (0, 1)$.

Proof. Let any ν_X, ν_Y probability measures on $\mathcal{X}$. Then, there exist random variables X_0, Y_0 independent of $\{\omega_n\}_{n\in\mathbb{N}}$ with

$$\mathcal{W}_1(\nu_X, \nu_Y) = \mathbb{E}[\|X_0 - Y_0\|_2] \leq \mathbb{E}\left[\|X_0 - Y_0\|_2^2\right]^{1/2} < \infty \tag{31}$$

where the inequality follows from Hölder's inequality. Then, for $X_1 = F(X_0, \omega)$ and $Y_1 = F(Y_0, \omega)$ that share the same randomness ω, we have by (20) that

$$\mathbb{E}\left[\|X_1 - Y_1\|_2^2 \mid X_0, Y_0\right] \leq (1 - 2\gamma\lambda + \gamma^2 NL)\|X_0 - Y_0\|_2^2 \tag{32}$$

Integrating (34) with respect to the law of (X_0, Y_0) given in (33), we get that

$$\mathbb{E}\left[\|X_1 - Y_1\|_2^2\right] \leq (1 - 2\gamma\lambda + \gamma^2 NL)\ \mathbb{E}\left[\|X_0 - Y_0\|_2^2\right] \tag{33}$$

Now, by the definition of the Wasserstein distance and invoking Hölder's inequality and (35), we have that

$$\begin{aligned} \mathcal{W}_1^2(\nu_X P, \nu_Y P) &\leq \mathbb{E}[\|X_1 - Y_1\|_2]^2 \leq \mathbb{E}\left[\|X_1 - Y_1\|_2^2\right] \\ &\leq (1 - 2\gamma\lambda + \gamma^2 NL)\ \mathbb{E}\left[\|X_0 - Y_0\|_2^2\right] \end{aligned} \tag{34}$$

Therefore, using induction, we get for any $k \geq 1$

$$\begin{aligned} \mathcal{W}_1^2(\nu_X P^{(k)}, \nu_Y P^{(k)}) &\leq \mathbb{E}\left[\|X_k - Y_k\|_2^2\right] \\ &\leq (1 - 2\gamma\lambda + \gamma^2 NL)^k\ \mathbb{E}\left[\|X_0 - Y_0\|_2^2\right] \end{aligned} \tag{35}$$

Now, setting $\nu_X \equiv \delta_x$, i.e., the Dirac measure at state x, and $\nu_Y \equiv \mu$, i.e., the unique invariant measure μ, and noting that $\mu = \mu P^{(k)}$ due to the invariance property of μ, (37) gives

$$\mathcal{W}_1^2(P_x^{(k)}, \mu) \leq (1 - 2\gamma\lambda + \gamma^2 NL)^k\ \mathbb{E}\left[\|X_0 - Y_0\|_2^2\right] \tag{36}$$

We thus get

$$\mathcal{W}_1(P_x^{(k)}, \mu) \leq (1 - 2\gamma\lambda + \gamma^2 NL)^{k/2}\, \mathbb{E}\left[\|X_0 - Y_0\|_2^2\right]^{1/2} \tag{37}$$

and the result follows. ■

Now, building on the convergence of the law of the process $\{X_n\}_{n\in\mathbb{N}}$, we can show the convergence of the mean occupation measure at rate $\mathcal{O}(1/n)$. Specifically, we have:

Theorem 5. *Let the Markov process $(X_n)_{n\in\mathbb{N}}$ generated by the (SPG) for $\gamma < 2\lambda/NL$, and $x \in \mathcal{X}$. Then, the mean occupation measure $\mu_n(x, \cdot)$ per (23), converges to the invariant measure μ at a rate*

$$\mathcal{W}_1(\mu_n(x, \cdot), \mu) = \mathcal{O}(1/n). \tag{38}$$

Proof. Let $\pi_k : \mathcal{X} \times \mathcal{X} \to [0, 1]$ be the optimal coupling between $P_x^{(k)}$ and μ, i.e.,

$$\mathcal{W}_1(P_x^{(k)}, \mu) = \int \|x - y\|_2 d\pi_k(x, y) \tag{39}$$

Then, setting $\Pi_n \equiv \sum_{k=0}^{n-1} \pi_k/n$, it is easy to see that Π_n is a coupling between $\mu_n(x, \cdot)$ and μ, since for any $A \in \mathcal{B}(\mathcal{X})$

$$\Pi_n(A \times \mathcal{X}) = \frac{1}{n}\sum_{k=0}^{n-1} \pi_k(A \times \mathcal{X}) = \frac{1}{n}\sum_{k=0}^{n-1} P_x^{(k)}(A) = \mu_n(x, A) \tag{40}$$

and

$$\Pi_n(\mathcal{X} \times A) = \frac{1}{n}\sum_{k=0}^{n-1} \pi_k(\mathcal{X} \times A) = \frac{1}{n}\sum_{k=0}^{n-1} \mu(A) = \mu(A) \tag{41}$$

Thus, for $\rho \equiv \left(1 - 2\gamma\lambda + \gamma^2 NL\right) \in (0, 1)$, we readily get:

$$\begin{aligned} \mathcal{W}_1(\mu_n(x, \cdot), \mu) &\leq \int \|x - y\|_2 d\Pi_n(x, y) \\ &= \frac{1}{n}\sum_{k=0}^{n-1} \mathcal{W}_1(P_x^{(k)}, \mu) \\ &\leq \frac{1}{n}\sum_{k=0}^{n-1} \rho^{k/2}\ \mathbb{E}\left[\|X_0 - Y_0\|_2^2\right]^{1/2} \\ &\leq \frac{1}{n}\ \mathbb{E}\left[\|X_0 - Y_0\|_2^2\right]^{1/2} \frac{1}{1 - \sqrt{\rho}} \end{aligned} \tag{42}$$

where the first inequality holds due to the definition of the Wasserstein-1 distance and the second one due to Theorem 4. Thus, the result follows. ■

The difference in the convergence rates of the law $P_x^{(n)}$ and the mean occupation measure $\mu_n(x, \cdot)$ is rather natural. While $P_x^{(n)}$ converges geometrically to μ, the Cesàro average $\sum_{k=0}^{n-1} P_x^{(k)}/n$ accumulates early "transient" terms, so their overall impact shrinks only on the order of $1/n$, introducing additional randomness and variability. In particular, the contribution from initial steps is spread out but not rapidly suppressed, leading to the slower $1/n$ convergence rate for the mean occupation measure. Moreover, it is important to highlight the following trade-off: decreasing the step-size γ increases the concentration of iterates within the ball $\mathbb{B}_\delta$, but results in slower convergence, as indicated by the constant ρ.

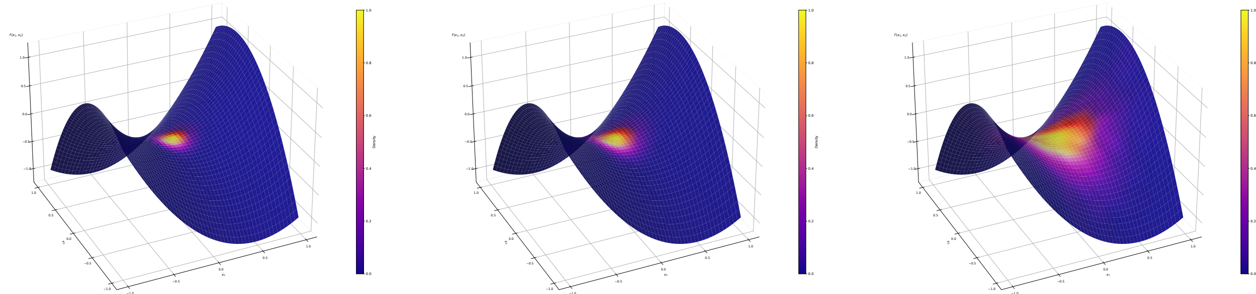
V. Numerical Experiments

In this section, we provide numerical simulations to illustrate and explore our theoretical findings. To this end, we consider a simple, yet illustrative, strongly monotone two-player min-max game defined by $f : [0, 1] \times [0, 1] \to \mathbb{R}$ with

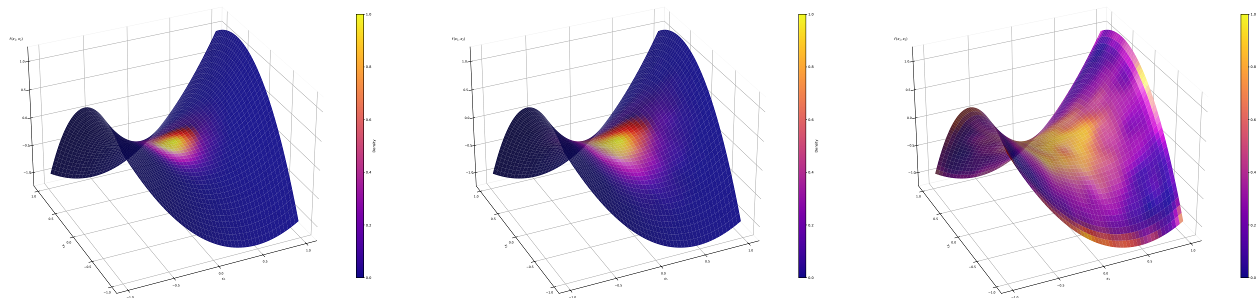
$$f(x_1, x_2) = x_1^2 - x_2^2 + x_1 x_2 \tag{43}$$

To be more precise, the payoff functions of the two players are given by $u_1(x_1, x_2) = f(x_1, x_2) = -u_2(x_1, x_2)$, and $x^* = (0, 0)$ is the unique Nash equilibrium point.

Fig. 1 demonstrates the behavior of (SPG) under varying step sizes and noise levels for the min-max game defined by the function $f(x_1, x_2)$. Specifically, we consider step-sizes $\gamma \in \{0.1, 0.2, 0.5\}$, and stochastic feedback of the form $V(x, \omega) = v(x) + \sigma\omega$, where $\omega \sim N(0, \mathbf{I}_2)$ for $\sigma \in \{0.5, 1\}$. For each (γ, σ) configuration, we conducted 10^5 separate trials, each for 10^3 steps, where the initial state for each run was drawn uniformly at random in $[-1, 1]^2$. Each surface



(a) $\sigma = 0.5$ and $\gamma = 0.1$ (left), $\gamma = 0.2$ (middle), $\gamma = 0.5$ (right)



(b) $\sigma = 1$ and $\gamma = 0.1$ (left), $\gamma = 0.2$ (middle), $\gamma = 0.5$ (right)

Fig. 1: Visualization of the long-run occupancy measure of (SPG) on the min-max game with loss-gain function $f(x_1, x_2) = x_1^2 - x_2^2 + x_1 x_2$. In each plot, 10^5 instances of (SPG) were run for 10^3 iterations, and the heatmap represents the empirical frequency of observing the last iterate of (SPG) at a certain point, overlain with the loss-gain landscape of the game.

represents the landscape of f, while the color overlay visualizes the empirical density of the *final iterates* across the 10^5 different trials. Warmer (bright/yellow) regions indicate higher concentration of final iterates, whereas cooler (purple/blue) regions correspond to lower probability of ending in those regions, as the colorbar on the side suggests. The first row corresponds to $\sigma = 0.5$, while the second to $\sigma = 1$; the three columns in each row correspond to (i) $\gamma = 0.1$, (ii) $\gamma = 0.2$, and (iii) $\gamma = 0.5$.

We observe that smaller step-size and noise parameters lead to tighter concentration around $x^* = (0, 0)$, whereas, as the noise increases, the final distribution becomes more spread out. This is expected since the level of the noise directly affects the variability of the iterates, reflecting the growing influence of stochastic perturbations on the dynamics. Additionally, in each column, we can see that the heatmap on the bottom is broader and less peaked than the corresponding heatmap on the top, due to the extra randomness that increases the spread in the final distribution of points. To conclude, a larger step-size and larger noise level each promote greater variance in the final positions reached by (SPG), whereas a smaller step-size and smaller noise lead to a tighter concentration of the distribution around the equilibrium.

VI. Concluding remarks

In this work, we studied the long-run behavior of the *stochastic projected gradient* (SPG) iterates in strongly monotone games with constant step-size. Specifically, we showed that the induced Markov process $(X_n)_{n\in\mathbb{N}}$ admits a unique invariant distribution, to which both the law of the process and the mean occupation measure converge in Wasserstein distance, irrespective of the initial state. Finally, our experiments illustrate how the concentration of the limiting distribution depends on the step-size and noise level.

References

[1] L. Pinto, J. Davidson, R. Sukthankar, and A. Gupta, "Robust adversarial reinforcement learning," in *ICML '17: Proceedings of the 34th International Conference on Machine Learning*, 2017.

[2] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu *et al.*, "Generative adversarial nets," in *NIPS '14: Proceedings of the 28th International Conference on Neural Information Processing Systems*, 2014.

[3] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," https://arxiv.org/abs/1701.07875, 2017.

[4] Y. Shoham and K. Leyton-Brown, *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press, 2008.

[5] M. Chiang, P. Hande, T. Lan, and C. W. Tan, "Power control in wireless cellular networks," *Foundations and Trends in Networking*, vol. 2, no. 4, pp. 381–533, 2007.

[6] Z. Zhou, P. Mertikopoulos, A. L. Moustakas, N. Bambos *et al.*, "Robust power management via learning and game design," *Operations Research*, vol. 69, no. 1, pp. 331–345, January 2021.

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

[8] S. Shalev-Shwartz, "Online learning and online convex optimization," *Foundations and Trends in Machine Learning*, vol. 4, no. 2, 2011.

[9] P. Mertikopoulos, Y.-P. Hsieh, and V. Cevher, "A unified stochastic approximation framework for learning in games," *Math. Program.*, vol. 203, no. 1–2, p. 559–609, Aug. 2023. [Online]. Available: https://doi.org/10.1007/s10107-023-02001-y

[10] L. Flokas, E. V. Vlatakis-Gkaragkounis, and G. Piliouras, “Poincaré recurrence, cycles and spurious equilibria in gradient-descent-ascent for non-convex non-concave zero-sum games,” in *NeurIPS ’19: Proceedings of the 33rd International Conference on Neural Information Processing Systems*, 2019.

[11] P. Mertikopoulos, C. H. Papadimitriou, and G. Piliouras, “Cycles in adversarial regularized learning,” in *SODA ’18: Proceedings of the 29th annual ACM-SIAM Symposium on Discrete Algorithms*, 2018.

[12] M. Posch, “Cycling in a stochastic learning algorithm for normal form games,” *Journal of Evolutionary Economics*, vol. 7, pp. 193–207, 1997.

[13] J. B. Rosen, “Existence and uniqueness of equilibrium points for concave N-person games,” *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.

[14] U. V. Shanbhag, *Stochastic Variational Inequality Problems: Applications, Analysis, and Algorithms*, ch. Chapter 5, pp. 71–107.

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

[16] M. Bravo, D. S. Leslie, and P. Mertikopoulos, “Bandit learning in concave N-person games,” in *NeurIPS ’18: Proceedings of the 32nd International Conference of Neural Information Processing Systems*, 2018.

[17] G. M. Korpelevich, “The extragradient method for finding saddle points and other problems,” *Èkonom. i Mat. Metody*, vol. 12, pp. 747–756, 1976.

[18] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, ser. Springer Series in Operations Research. Springer, 2003.

[19] P. Tseng, “On linear convergence of iterative methods for the variational inequality problem,” *Journal of Computational and Applied Mathematics*, vol. 60, no. 1-2, pp. 237–252, June 1995.

[20] A. Juditsky, A. S. Nemirovski, and C. Tauvel, “Solving variational inequalities with stochastic mirror-prox algorithm,” *Stochastic Systems*, vol. 1, no. 1, pp. 17–58, 2011.

[21] E. Gorbunov, H. Berard, G. Gidel, and N. Loizou, “Stochastic extragradient: General analysis and improved rates,” in *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, ser. Proceedings of Machine Learning Research, G. Camps-Valls, F. J. R. Ruiz, and I. Valera, Eds., vol. 151. PMLR, 28–30 Mar 2022, pp. 7865–7901.

[22] A. S. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro, “Robust stochastic approximation approach to stochastic programming,” *SIAM Journal on Optimization*, vol. 19, no. 4, pp. 1574–1609, 2009.

[23] A. Kannan and U. V. Shanbhag, “Optimal stochastic extragradient schemes for pseudomonotone stochastic variational inequality problems and their variants,” *Computational Optimization and Applications*, vol. 74, pp. 779 – 820, 2014.

[24] J. Yang, N. Kiyavash, and N. He, “Global convergence and variance reduction for a class of nonconvex-nonconcave minimax problems,” in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan *et al.*, Eds., vol. 33. Curran Associates, Inc., 2020, pp. 1153–1165.

[25] N. Loizou, H. Berard, G. Gidel, I. Mitliagkas *et al.*, “Stochastic gradient descent-ascent and consensus optimization for smooth games: convergence analysis under expected co-coercivity,” ser. NIPS ’21. Red Hook, NY, USA: Curran Associates Inc., 2021.

[26] P. Mertikopoulos, B. Lecouat, H. Zenati, C.-S. Foo *et al.*, “Optimistic mirror descent in saddle-point problems: Going the extra (gradient) mile,” in *ICLR ’19: Proceedings of the 2019 International Conference on Learning Representations*, 2019.

[27] Y.-G. Hsieh, F. Iutzeler, J. Malick, and P. Mertikopoulos, “Explore aggressively, update conservatively: Stochastic extragradient methods with variable stepsize scaling,” in *NeurIPS ’20: Proceedings of the 34th International Conference on Neural Information Processing Systems*, 2020.

[28] W. Azizian, F. Iutzeler, J. Malick, and P. Mertikopoulos, “The last-iterate convergence rate of optimistic mirror descent in stochastic variational inequalities,” in *COLT ’21: Proceedings of the 34th Annual Conference on Learning Theory*, 2021.

[29] Y.-G. Hsieh, K. Antonakopoulos, and P. Mertikopoulos, “Adaptive learning in continuous games: Optimal regret bounds and convergence to Nash equilibrium,” in *COLT ’21: Proceedings of the 34th Annual Conference on Learning Theory*, 2021.

[30] Y.-G. Hsieh, K. Antonakopoulos, V. Cevher, and P. Mertikopoulos, “No-regret learning in games with noisy feedback: Faster rates and adaptivity via learning rate separation,” in *NeurIPS ’22: Proceedings of the 36th International Conference on Neural Information Processing Systems*, 2022.

[31] E. Vasileios Vlatakis-Gkaragkounis, A. Giannou, Y. Chen, and Q. Xie, “Stochastic methods in variational inequalities: Ergodicity, bias and refinements,” in *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, ser. Proceedings of Machine Learning Research, S. Dasgupta, S. Mandt, and Y. Li, Eds., vol. 238. PMLR, 02–04 May 2024, pp. 4123–4131.

[32] S. Meyn, R. L. Tweedie, and P. W. Glynn, *Markov Chains and Stochastic Stability*, 2nd ed., ser. Cambridge Mathematical Library. Cambridge University Press, 2009.

[33] L. Yu, K. Balasubramanian, S. Volgushev, and M. A. Erdogdu, “An analysis of constant step size sgd in the non-convex regime: asymptotic normality and bias,” in *Proceedings of the 35th International Conference on Neural Information Processing Systems*, ser. NIPS ’21. Red Hook, NY, USA: Curran Associates Inc., 2021.

[34] A. Dieuleveut, A. Durmus, and F. R. Bach, “Bridging the gap between constant step size stochastic gradient descent and markov chains,” *The Annals of Statistics*, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:4605591

[35] B. T. Polyak and A. Juditsky, “Acceleration of stochastic approximation by averaging,” *SIAM Journal on Control and Optimization*, vol. 30, no. 4, pp. 838–855, July 1992.

[36] D. Ruppert, “Efficient estimations from a slowly convergent robbins-monro process,” 02 1988.

[37] A. Anastasiou, K. Balasubramanian, and M. A. Erdogdu, “Normal approximation for stochastic gradient descent via non-asymptotic rates of martingale clt,” in *Proceedings of the Thirty-Second Conference on Learning Theory*, ser. Proceedings of Machine Learning Research, A. Beygelzimer and D. Hsu, Eds., vol. 99. PMLR, 25–28 Jun 2019, pp. 115–137.

[38] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton University Press, 1970.

[39] C. Villani, *Optimal Transport: Old and New*. Springer Berlin, Heidelberg, 2008.

[40] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, ser. Applied Optimization. Kluwer Academic Publishers, 2004, no. 87.

[41] O. Hernández-Lerma and J.-B. Lasserre, *Markov Chains and Invariant Probabilities*. Birkhäuser Basel, 2003.

[42] R. Douc, E. Moulines, P. Priouret, and P. Soulier, *Markov chains*, ser. Operation research and financial engineering. Springer, 2018.

[43] G. B. Folland, *Real Analysis*, 2nd ed. Wiley-Interscience, 1999.

[44] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite Dimensional Systems*, ser. London Mathematical Society Lecture Note Series. Cambridge University Press, 1996.

[45] M. Hairer, *Ergodic properties of Markov processes*. Lecture notes, 2006.

[46] ——, *Convergence of Markov processes*. Lecture notes, 2010.