Some applications of new results in stochastic approximation with discontinuous drifts

Quoc Le
Department of Mathematics and Applied Mathematical Sciences
University of Rhode Island
Rhode Island, USA
quocle@uri.edu

Nhu Nguyen
Department of Mathematics and Applied Mathematical Sciences
University of Rhode Island
Rhode Island, USA
nhu.nguyen@uri.edu

George Yin
Department of Mathematics
University of Connecticut
Conneticut, USA
gyin@uconn.edu

***Abstract*—The paper is concerned with stochastic approximation algorithms. Our main effort is focused on recently developed set-valued stochastic approximation methods. We begin with a brief introduction on stochastic approximation. Next, recent results are reviewed. Then the rest of the paper concentrates on applications of stochastic applications of set-valued problems.**

***Index Terms*—stochastic approximation, stochastic optimization, recursive algorithm, set-valued analysis, convergence, rate of convergence, stochastic inclusion.**

I. Introduction

SEVENTY-FIVE years have passed since the stochastic approximation (SA) methods were introduced by Robbins and Monro in their pioneering work [22], which was named RM algorithm later. During the years, significant progress has been made in the study of such stochastic recursive defined algorithms. The development of SA has been always tied-up with real applications, especially various forms of stochastic optimization problems, as well as control, signal processing, image processing, pattern classifications, and a wide range of applications in related fields. The original motivation of Robbins and Monro stems from the problem of finding roots of a continuous function $f(\cdot)$, where either the precise form of the function is not known, or it is too complicated to compute; the experimenter is able to take "noisy" measurements at desired values, however. A classical example is to find appropriate dosage level of a drug, provided only $f(x)$+noise is available, where x is the level of dosage and $f(x)$, assumed to be an increasing function, is the probability of success (leading to the recovery of the patient) at dosage level x. The classical Kiefer–Wolfowitz (KW) algorithm introduced by Kiefer and Wolfowitz the concerns the minimization of a real-valued function using only noisy functional measurements. In both RM and KW algorithms, the main concerns of the theoretical issues focus on analysis of iteratively defined stochastic processes and a wide variety of applications focus on the basic paradigm of stochastic difference equations. Much of the development has been originated from a wide range of applications in optimization, control theory, economic systems, signal processing, communication theory, learning, pattern classification, neural network, and many other related fields. Owing to its importance, stochastic approximation has had a long history and has drawn much attention in the past five decades. A number of monographs have been written; each of them has its own distinct features. To mention just a few, we cite the books of Albert and Gardner [1], Wasan [30], Tsypkin [29], Nevel'son and Khasminskii [20], Kushner and Clark [16], Benveniste, Métivier, and Priouret [6], Duflo [11], Solo and Kong [25], Chen and Zhu [9], and Kushner and Yin [17] among others.

The work of Quoc Le and Nhu Nguyen was supported in part by the National Science Foundation under the grant DMS-2407669; the work of G. Yin was supported in part by the National Science Foundation under the grant DMS-2404508.

In fact, stochastic approximation methods have been the subject of an enormous literature, both theoretical and applied for five decades. Due to the vast amount of literature accumulated, it is very difficult or virtually impossible to provide an exhaustive list of references on stochastic approximation. The development of the stochastic approximation methods can be naturally divided into several periods. To put things in the historical perspective, we mention some of the works on stochastic approximation in what follows. The early development around 1950's and 1960's used mainly basic probabilistic tools and traditional statistical assumptions (such as independent and identically distributed noise) together with certain restrictions on the functions (such as assuming $f(x)$ to be increasing for instance).

A. Basic Setup

This work aims to present some applications of very recent results on stochastic approximation and from new angles. Our main concerns are that the dynamics of the systems involved are not necessarily smooth or even continuous. In lieu of the set up of working with vector-valued iterations, we examine set-valued processes. Technically it becomes more challenging. In the study of stochastic approximation, we normally show the convergence of the algorithms by showing that suitably interpolated sequences converge to a deterministic ordinary differential equations (ODEs). The rates of convergence was studied by showing suitably scaled and centered sequences converge to stochastic differential equations. Although there were some early results on replacing the ODEs by differential inclusions when set-valued process were considered, little was known for the related rates of convergence. Our recent results [21] obtained stochastic differential inclusion limits so as to substantially generalized the existing literature.

Precisely, a stochastic approximation algorithm has the form

$$\mathbf{X}_{n+1} = \mathbf{X}_n + a_n \mathbf{b}_n (\mathbf{X}_n, \xi_n) + a_n \boldsymbol{\beta}_n, \tag{I.1}$$

where $\{a_n\}$'s are step sizes satisfying $a_n \in \mathbb{R}^+, a_n \to 0$, and $\sum_{n=1}^{\infty} a_n = \infty$, the sequence $\{\xi_n\}$ is a noise process, and $\{\boldsymbol{\beta}_n\}$ represents the bias. If the sequence $\mathbf{b}_n(\cdot, \xi_n)$ and its associated "noise averaging-out" function are continuous, the asymptotic properties of the algorithms have been well-understood, see e.g., [15], [16], [17], [18], [19] and references therein.

Motivated by optimization problems with non-differentiable loss functions and the search for zero points of set-valued mappings perturbed by random noise, [21] initiates the study of stochastic approximation in the context of discontinuous dynamics and set-valued mappings within a general and unified framework. Furthermore, new techniques are introduced for analyzing algorithms involving set-valued analysis and stochastic differential inclusions. By allowing the functions $\mathbf{b}_n(\cdot, \cdot)$ to be discontinuous and belong to a set-valued mapping, the work in [21] opens new avenues for applying stochastic approximation beyond optimization problems to a broader range of fields. As an example of practical algorithms of (I.1) in this discontinuous setting, consider the sign-error algorithm [31] frequently used in adaptive filtering, which uses the sign operator to reduce the computational complexity:

$$\theta_{n+1} = \theta_n + a_n \varphi_n \text{sign}(y_n - \varphi_n^\top \theta_n).$$

Another example is the stochastic version of sub-gradient descent algorithm for stochastic optimization problem with non-differential loss function often used in support vector machine (SVM) classification problem:

$$\mathbf{w}_{n+1} = \mathbf{w}_n - a_n \lambda \mathbf{w}_n + a_n g_n(\mathbf{w}_n, \mathbf{x}_n, y_n), \tag{I.2}$$

where $g_n(\mathbf{w}_n, \mathbf{x}_n, y_n) \in \partial\big(\text{-}\max\{0, 1 - y_n \mathbf{w}_n^\top \mathbf{x}_n\}\big)$, i.e.,

$$g_n(\mathbf{w}_n, \mathbf{x}_n, y_n) \in \begin{cases} \{\mathbf{0}\} \text{ if } y_n \mathbf{w}_n^\top \mathbf{x}_n > 1, \\ \overline{\text{co}}\ \{\mathbf{0}, y_n \mathbf{x}_n\} \text{ if } y_n \mathbf{w}_n^\top \mathbf{x}_n = 1, \\ \{y_n \mathbf{x}_n\} \text{ if } y_n \mathbf{w}_n^\top \mathbf{x}_n < 1. \end{cases}$$

This work will present applications of our new results in our paper [21] with focuses on multistage decision making problem and game theory. The "multistage decision-making model" examined in this work involves a sequential decision process where each choice influences the available options and possible outcomes in future stages. This model holds significant importance in economics, drawing motivation in part from Smale's approach to the prisoner's dilemma [24], Blackwell's approachability theory [7], [26], fictitious play [8], [23], and stochastic fictitious play [3], [12], [13]. Unlike many existing studies on multistage decision-making (see, e.g., [5] and the references therein), our framework allows the decision-maker to observe outcomes only partially and under noise. Additionally, we characterize the limiting processes as solutions rather than perturbed solutions of the corresponding differential inclusion, while also deriving results on convergence rates and robustness. This results can be further extended to accommodate alternative optimality criteria, such as overtaking and bias, as well as to analyze more complex systems, including switching dynamical systems.

From another perspective, we investigate applications in fictitious play, where each player, at each stage of a repeated game, selects a move that best responds to the past frequency of the opponent's actions. In particular, we analyze a two-player game in which Player 1 has limited information—either they do not recall their past moves, are unaware of the opponent's action set, or are not informed of the opponent's chosen moves. We will generalize existing settings by assuming that players observes environments under noise perturbations. Then, we investigate its important asymptotic properties.

The organization of the paper is as follows. Section II present briefly the theoretical results in our previous work [21], which we will need here. Section III-A studies multistage decision-making problem with partial observations. Finally, a two-players game problem is considered in Section III-B.

II. Theoretical results

This section is devoted to presenting briefly theoretical results in our previous paper [21]. Let $t_0 = 0$, $t_n := \sum_{i=0}^{n-1} a_i$ for $n \geq 1$, and

$$m(t) := \begin{cases} \max\{n : t_n \leq t\} & \text{if } t \geq 0, \\ 0 & \text{if } t < 0. \end{cases}$$

We define the piecewise linear interpolation $\mathbf{X}^0(t)$ of $\mathbf{X}_n$ with interpolation intervals $\{a_n\}$ by $\mathbf{X}^0(t_n) := \mathbf{X}_n$, and

$$\mathbf{X}^0(t) := \frac{t_{n+1} - t}{a_n}\mathbf{X}_n + \frac{t - t_n}{a_n}\mathbf{X}_{n+1} \text{ in } (t_n, t_{n+1}),$$

and the shift sequence $\mathbf{X}^n(\cdot)$ on $(-\infty, \infty)$ is defined as

$$\mathbf{X}^n(t) := \begin{cases} \mathbf{X}^0(t + t_n) & \text{if } t \geq -t_n, \\ \mathbf{X}_0 & \text{if } t \leq -t_n. \end{cases}$$

We need the following conditions.

Assumption 1. There is a set-valued mapping $G : \mathbb{R}^d \to 2^{\mathbb{R}^d}$ satisfying:

(i) There is a finite ball $B_G \subset \mathbb{R}^d$ such that $G(\mathbf{x}) \subset B_G$ for all $\mathbf{x}$, which means $G(\cdot)$ has non-empty, compact, convex values, and all values are contained in a finite common ball;

(ii) The graph of G, $\text{Graph}(G) := \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d \times \mathbb{R}^d : \mathbf{y} \in G(\mathbf{x})\}$, is closed;

(iii) There exists a sequence of positive real-valued functions $\{m_n(\mathbf{x}, \xi)\}$ continuous in $\mathbf{x}$ and uniformly in ξ such that

$$\mathbf{b}_n(\mathbf{x}, \xi) \in G(\mathbf{x}) + m_n(\mathbf{x}, \xi)\bar{B} \quad \text{for all } n, \mathbf{x}, \xi,$$

and that for some $T > 0$, each $\varepsilon > 0$, and each $\mathbf{x}$,

$$\lim_{n\to\infty} \mathbb{P}\left\{ \sup_{j\geq n} \max_{t\leq T} \left| \sum_{i=m(jT)}^{m(jT+t)-1} a_i (m_i(\mathbf{x}, \xi_i) + \beta_i) \right| \geq \varepsilon \right\} = 0.$$

Under the Assumption 1, by using [21, Theorems 2.2 and 2.4], we obtain the following result.

Theorem II.1. *Consider algorithm (I.1) with Assumption 1. Suppose that $\{\mathbf{X}_n\}$ is bounded w.p.1.*

- *Then there is a null set Ω_0 such that $\forall \omega \notin \Omega_0, \{\mathbf{X}^n(\cdot)\}$ is bounded and equicontinuous on bounded intervals.*
- *Let $\mathbf{X}(\cdot)$ be the limit of a convergent subsequence of $\{\mathbf{X}^n(\cdot)\}$. Then $\mathbf{X}(\cdot)$ is a solution of the differential inclusion*

$$\dot{\mathbf{X}}(t) \in G(\mathbf{X}(t)) \tag{II.1}$$

- *The limit set of $\mathbf{X}(\cdot)$ is internally chain transitive (with respect to (II.1)) and the limit points of $\{\mathbf{X}_n\}$ are contained in $\mathcal{R}$, the set of chain-recurrent points of (II.1).*
- *Moreover, let Λ be a locally asymptotically stable set (in the sense of Lyapunov) of all solutions of (II.1) and $DA(\Lambda)$ be its domain of attraction. If $\{\mathbf{X}_n\}$ visits the compact subset of $DA(\Lambda)$ infinitely often with probability 1 (resp., with probability $\geq \rho$), then $\mathbf{X}_n \rightarrow \Lambda$ when $n \rightarrow \infty$ with probability 1 (resp., with probability $\geq \rho$).*
- *Assume further that there is a unique $\mathbf{x}^*$ such that $\mathbf{0} \in G(\mathbf{x}^*)$; and that there exists a $\mathcal{U}$-generalized Lyapunov function $V: \mathbb{R}^d \rightarrow \mathbb{R}_+$ such that the sublevel sets $\{x \in \mathbb{R}^n : V(\mathbf{x}) \leq l\}$ are compact for every $l > 0$ and the $\mathcal{U}$-generalized derivative $\dot{V}_{\mathcal{U}}^{G^*}(\mathbf{x})$ satisfies the "decay condition" in the sense of [21, Assumption (GS)] with $G^*(\mathbf{x}) = G(\mathbf{x}+\mathbf{x}^*)$. Then, $\mathbf{X}_n \rightarrow \mathbf{x}^*$ w.p.1.*

Next, consider the case that the noise cannot be averaged out, that means Assumption 1(iii) does not hold.

Definition II.1. For a set $A \subset \mathbb{R}^d$, an ε-neighborhood of A denoted by $N_\varepsilon(A)$ is defined as $N_\varepsilon(A) = \left\{\mathbf{x} \in \mathbb{R}^d : \operatorname{distance}(\mathbf{x}, A) \leq \varepsilon\right\}$, where $\operatorname{distance}(\mathbf{x}, A) := \inf_{\mathbf{y} \in A} |\mathbf{x}-\mathbf{y}|$.

Utilizing [21, Theorem 2.6] under Assumption 1, we get the following theorem.

Theorem II.2. *Consider algorithm (I.1) with Assumption 1 (i)-(ii) and let $\eta := \limsup_{n \rightarrow \infty} \|m_n\| + |\beta_n| > 0$. Assume that $\{\mathbf{X}_n\}$ is bounded w.p. 1.*

- *Then, there is a null set Ω_0 such that $\forall \omega \notin \Omega_0, \{\mathbf{X}^n(\cdot)\}$ is bounded and equicontinuous. If we let $\mathbf{X}(\cdot)$ be the limit of a convergent subsequence of $\{\mathbf{X}^n(\cdot)\}$, then $\mathbf{X}(\cdot)$ is a solution of the differential inclusion $\dot{\mathbf{X}}(t) \in N_{2\eta}(G(\mathbf{X}(t)))$.*
- *There exists a (deterministic) positive function $\phi(\cdot): [0, \infty) \rightarrow [0, \infty)$ depending on $\limsup_n |\mathbf{X}_n|$ (resp., the projection space A) such that $\lim_{t \rightarrow 0} \phi(t) = \phi(0) = 0$ and $\limsup_{n \rightarrow \infty} \operatorname{distance}(\mathbf{X}_n, \mathcal{R}) \leq \phi(\eta)$, where $\mathcal{R}$ is the set of chain recurrent points of differential inclusion $\dot{\mathbf{X}}(t) \in G(\mathbf{X}(t))$*

Now, to study the rates of convergence, we consider the following simple algorithm with additive noise

$$\mathbf{X}_{n+1} = \mathbf{X}_n + a_n \mathbf{b}_n(\mathbf{X}_n) + a_n \boldsymbol{\beta}_n, \quad \mathbf{b}_n(\mathbf{X}_n) \in G(\mathbf{X}_n). \tag{II.2}$$

The following condition is needed for investigating the rate of convergence.

Assumption 2.
(i) The sequence of step sizes $\{a_n\}_{n \geq 0}$ satisfies $0 < a_n \rightarrow 0$ as $n \rightarrow \infty$ and $(a_n / a_{n+1})^{1/2} = 1 + \varepsilon_n$ where (a) $\varepsilon_n = \frac{1}{2n} + o(\varepsilon_n)$ if $a_n = 1/n$, or (b) $\varepsilon_n = o(a_n)$.
(ii) There is a limit point $\mathbf{x}^*$ satisfying the following conditions: (a) $\mathbf{X}_n \rightarrow \mathbf{x}^*$ w.p. 1 and $\overline{\mathbf{h}}(\mathbf{x}^*) + G(\mathbf{x}^*) = \{\mathbf{0}\}$; (b) $\left\{(\mathbf{X}_n - \mathbf{x}^*) / \sqrt{a_n}\right\}$ is tight.
(iii) The functions $\mathbf{h}(\cdot, \cdot)$ and $\mathbf{h}_{\mathbf{x}}(\cdot, \cdot)$ (gradient with respect to $\mathbf{x}$) are continuous in ($\mathbf{x}, \xi$) and bounded on bounded $\mathbf{x}$-sets. The second partial derivative (with respect to $\mathbf{x}$)$\mathbf{h}_{\mathbf{xx}}(\cdot, \xi)$ exists and is bounded uniformly in ξ, and $\mathbf{h}_{\mathbf{xx}}(\cdot, \xi)$ is continuous in a neighborhood of $\mathbf{x}^*$. The $\{\xi_n\}$ is a sequence of uniformly bounded and satisfies certain mixing conditions as in [21, Assumption R], for a detailed discussion and technical formulation on these mixing conditions, see [21] and references therein.
(iv) The set-valued mapping $G(\cdot)$ has non-empty, convex, and compact values, which are contained in a finite common ball such that $\mathbf{b}_n(\mathbf{x}) \in G(\mathbf{x}) \forall n$. Moreover, there is a continuous and positively homogeneous set-valued mapping T, whose values are non-empty, convex, compact, and contained in a finite common ball such that G is outer T-differentiable at $\mathbf{x}^*$ (see [21, Section A. 5] for these concepts).

We state [21, Theorem 3.1] for the rate of convergence of Algorithm (II.2).

Theorem II.3. *Consider algorithm (II.2) and assume Assumption 2 holds. Let $\{\mathbf{U}^n(\cdot)\}$ be the shift sequences of functions of the piecewise-constant interpolation generated by the normalized sequence $\frac{\mathbf{X}_n - \mathbf{x}^*}{\sqrt{a_n}}$. Then $\{\mathbf{U}^n(\cdot)\}$ converges weakly to the solutions of the stochastic differential inclusion*

$$d\mathbf{U}(t) \in [A\mathbf{U}(t) + T(\mathbf{U}(t))]dt + \Sigma_1^{1/2} d\overline{\mathbf{W}}(t)$$

if $(\mathbf{R})$(i)(a) *holds, and*

$$d\mathbf{U}(t) \in [(A + I/2)\mathbf{U}(t) + T(\mathbf{U}(t))]dt + \Sigma_1^{1/2} d\overline{\mathbf{W}}t\Big)$$

if $(\mathbf{R})$(i)(b) *holds. Where $\overline{\mathbf{W}}(t)$ is a d-dimensional standard Brownian motion.*

III. Applications

We demonstrate here the applicability of the proposed framework and the results presented in Section II. A natural application of (I.1) is for optimization of non-smooth loss functions, which are common in machine learning, as discussed in [21]. Here, however, we go beyond standard optimization problems and focus on applications in multistage decision-making under partial observations and two-player games in game theory. The applications presented below highlight two main contributions. First, we show that the original conditions can be relaxed. Second, we derive sharper theoretical results.

A. Multistage Decision Making with Partial Observations

Formulation. Let $\mathcal{E}$ and $\mathcal{B}$ be measurable spaces representing the action and state spaces, respectively. Consider a convex and compact set $\mathcal{O} \subset \mathbb{R}^d$ as the outcome space. At each discrete time step $n = 1, 2, \ldots$, a decision maker selects an action e_n from $\mathcal{E}$ and observes the outcome $M(e_n, b_n)$, where $M : \mathcal{E} \times \mathcal{B} \to \mathcal{O}$ is a measurable function. However, in many practical scenarios, the outcome is not fully observable; instead, it is subject to noise. Consequently, the decision maker does not have direct access to $M(e_n, b_n)$ but only to a noisy observation $\widetilde{M}(e_n, b_n, \xi_n)$, where ξ_n represents the noise.

We analyze a multistage decision-making model with partial observations under the following framework. The sequences $\{(e_n, b_n)\}_{n\geq 0}$ and $\{\xi_n\}_{n\geq 0}$ form random processes defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and they are adapted to the filtration $\{\mathcal{F}_n\}$. The noise sequence $\{\xi_n\}$ satisfies the condition that for some $T > 0$, any $\varepsilon > 0$, and any $(e, b) \in \mathcal{E} \times \mathcal{B}$,

$$\lim_{n\to\infty} \mathbb{P}\left\{\sup_{j\geq n} \max_{t\leq T} \left| \sum_{i=m(t)}^{m(jT+t)-1} \frac{1}{i} \Delta_i M \right| \geq \varepsilon \right\} = 0, \quad \text{(III.1)}$$

where $m(t) := \max\left\{n \in \mathbb{N} : \sum_{i=1}^{n} \frac{1}{i} \leq t\right\}$ and $\Delta_i M = M(e, b) - \widetilde{M}(e, b; \xi_i)$. The decision-making process is independent of the underlying environment given past information $\{(e_1, b_1), \ldots, (e_n, b_n)\}$, which can be expressed as $\mathbb{P}((e_{n+1}, b_{n+1}) \in de \times db \mid \mathcal{F}_n) = \mathbb{P}(e_{n+1} \in de \mid \mathcal{F}_n)\mathbb{P}(b_{n+1} \in db \mid \mathcal{F}_n)$. Rather than keeping track of individual observations, the decision maker maintains only the cumulative average of past (partially observed) outcomes,

$$\mathbf{X}_n = \frac{1}{n} \sum_{i=1}^{n} \widetilde{M}(e_i, b_i; \xi_i). \quad \text{(III.2)}$$

Decisions for future actions are based on this running average, following the probability distribution $\mathbb{P}(e_{n+1} \in de \mid \mathcal{F}_n) = Q_{\mathbf{x}_n}(de)$, where $Q_{\mathbf{x}}(\cdot)$ is a probability measure on $\mathcal{E}$ with finite second-moment for each $\mathbf{x} \in \mathcal{O}$. Moreover, for any measurable set $A \subset \mathcal{E}$, the function $\mathbf{x} \mapsto Q_{\mathbf{x}}(A)$ is measurable. The collection $Q = \{Q_{\mathbf{x}} : \mathbf{x} \in \mathcal{O}\}$ defines the strategy employed by the decision maker. Blackwell's approachability theory plays an important role in economics. We first make an application of results presented in Section II to investigate Blackwell's approachability.

Definition III.1. [Blackwell's approachability] A set $E \subset \mathcal{O}$ is said to be approachable if there exists a strategy Q such that $\mathbf{X}_n \to E$ w.p.1.

Let $\mathbf{X}^n(t)$ be the shift sequence of functions of the piecewise linear interpolation generated by the sequence $\{\mathbf{X}_n\}$ and the step-size sequence $\{\frac{1}{n}\}$.

For each $\mathbf{x} \in \mathcal{O}$, let

$$G_1(\mathbf{x}) = \left\{ \int_{\mathcal{E}\times\mathcal{B}} M(e, b) Q_{\mathbf{x}}(de)\nu(db) : \nu \in \mathcal{P}(\mathcal{B}) \right\},$$

where $\mathcal{P}(\mathcal{B})$ is the set of probability measures over $\mathcal{B}$ having finite second-moment. Define $G(\mathbf{x}) = -\mathbf{x} + \overline{\text{co}} G_1(\Pi_{\mathcal{O}}(\mathbf{x}))$. The next two theorems will provide conditions that a set is approachable.

Theorem III.1. *The limit of any convergent subsequence of $\mathbf{X}^n(t)$ is a solution of the following differential inclusion w.p.1*

$$\dot{\mathbf{X}}(t) \in G(\mathbf{X}(t)). \quad \text{(III.3)}$$

Proof. From (III.2), we have that

$$\begin{aligned}
&(n+1)(\mathbf{X}_{n+1} - \mathbf{X}_n) \\
&= -\mathbf{X}_n + \widetilde{M}(e_{n+1}, b_{n+1}; \xi_{n+1}) \\
&= -\mathbf{X}_n + \int_{\mathcal{E}\times\mathcal{B}} M(e, b_{n+1}) Q_{\mathbf{x_n}}(de) \mathbb{P}(b_{n+1} \in db \mid \mathcal{F}_n) \\
&\quad + \Big(\int_{\mathcal{E}} M(e, b_{n+1}) Q_{\mathbf{x_n}}(de) \\
&\qquad - \int_{\mathcal{E}\times\mathcal{B}} M(e, b_{n+1}) Q_{\mathbf{x_n}}(de) \mathbb{P}(b_{n+1} \in db \mid \mathcal{F}_n) \Big) \\
&\quad + M(e_{n+1}, b_{n+1}) - \int_{\mathcal{E}} M(e, b_{n+1}) Q_{\mathbf{x_n}}(de) \\
&\quad + \widetilde{M}(e_{n+1}, b_{n+1}; \xi_{n+1}) - M(e_{n+1}, b_{n+1}).
\end{aligned} \quad \text{(III.4)}$$

We can check that

$$-\mathbf{X}_n + \int_{\mathcal{E}\times\mathcal{B}} M(e, b_{n+1}) Q_{\mathbf{x_n}}(de) \mathbb{P}(b_{n+1} \in db \mid \mathcal{F}_n) \in G(\mathbf{x}),$$

and the remaining terms in (III.4) satisfy Assumption 1. Therefore, by applying Theorem II.1, we complete the proof. ☐

Theorem III.2. *If there is a strategy Q such that E is a globally asymptotically stable set of differential inclusion (III.3), then E is approachable.*

Proof. Since E is a globally asymptotically stable set of differential inclusion (III.3), $\{\mathbf{X}_n\}$ visits the compact subset of its domain of attraction DA(E) infinitely often with probability 1. Therefore, by applying Theorem II.1, $\mathbf{X}_n \to E$ w.p.1, which means E is approachable. We complete the proof. ☐

Next, we investigate the robustness of Blackwell's approachability. Assume that noise-corrupted perturbations cannot be averaged out but is bounded by η and denoted by E_η an approachable set of the corresponding algorithm with this η. We expect that Blackwell's approachability is robust in the sense that when $\eta \to 0$, E_η converges to an approachable set E of the algorithm corresponding to $\eta = 0$. In the next theorem, we show that this property holds.

Theorem III.3. *If the "convergence to 0 " condition (III.1) is relaxed as $|M(e, b) - \widetilde{M}(e, b, \xi)| < \eta, \forall e, b, \xi$, w.p.1, then theorems III.1 and III.2 still hold with G in (III.3) being replaced by its neighbor with radius η. Moreover, if E_η is a globally asymptotically stable set of the corresponding (limit) differential inclusions (and thus, is an approachable set), then there is a (deterministic) non-decreasing function $\phi(\cdot)$ satisfying* $\lim_{t\to 0} \phi(t) = 0$ *such that* distance $(E_\eta, E) \leq \phi(\eta)$.

Proof. Since $|M(e, b) - \widetilde{M}(e, b, \xi)| < \eta, \forall e, b, \xi$, w.p.1, using proposition 2.2 in [21], we get that $\dot{\mathbf{X}}(t) \in \overline{\text{co}}(G(\mathbf{X}(t) +$

$2\varepsilon B) + \eta B) = \overline{\text{co}}(G(\mathbf{X}(t) + 2\varepsilon B)) + \eta B$. Hence, $\dot{\mathbf{X}}(t) \in \cap_{\varepsilon>0}\overline{\text{co}}G(\mathbf{X}(t) + \varepsilon B) + \eta B = G(\mathbf{X}(t)) + \eta B$, or

$$\dot{\mathbf{X}}(t) \in N_\eta(G(\mathbf{X}(t))). \tag{III.5}$$

By using Theorem II.2 there exists a (deterministic) positive non-decreasing function $\phi(\cdot)$ such that $\lim_{t\to 0}\phi(t) = 0$ and $\limsup_{n\to\infty} \text{distance}(\mathbf{X}_n, E) \leq \phi(\eta)$. Thus, $\text{distance}\,(E_\eta, E) \leq \phi(\eta)$. We complete the proof. □

Next, we assume that there are a policy Q and $\mathbf{x}^*$, which is a unique approachable point under policy Q. We aim to investigate the rate of this approaching. Let $\mathbf{Y}_n := \frac{\mathbf{X}_n - \mathbf{x}^*}{\sqrt{n}}$ be the normalized sequence, and define the piecewise constant interpolation $\mathbf{Y}^0(\cdot)$ of $\mathbf{Y}_n$ and its shifted process $\mathbf{Y}^n(\cdot)$ as

$$\mathbf{Y}^0(t) := \mathbf{Y}_n \text{ if } t \in [t_n, t_{n+1}); \text{ and } \mathbf{Y}^n := \mathbf{Y}^0(t_n + t), t \geq 0.$$

Theorem III.4. *If there exists a strategy Q and $E = \{\mathbf{x}^*\}$ is a unique approachable set, then under Assumption 2, the limit process of a convergent subsequence of shifted interpolated processes generated by normalized sequence $\frac{\mathbf{X}_n - \mathbf{x}^*}{\sqrt{n}}$ converges weakly to solutions of a stochastic differential inclusion.*

Proof. By rewriting $\mathbf{X}_{n+1}$ as we did in (III.4), we can verify that the sequence $\mathbf{X}_n$ fit the framework of Theorem II.3. Thus, this theorem follows Theorem II.3 □

B. Two-Player Game

In this section, we consider a two-player game. Let S and V be two sets of moves, they can be infinite. At each stage n, player 1 choose $s_n \in S$ and player 2 choose $v_n \in V$. The payoff of that stage is $g(s_n, v_n)$, where $g : S \times V \to \mathcal{O}$ is a measurable function, $\mathcal{O} \subset \mathbb{R}^m$ is the outcome space. However, in practice, the exact outcome $g(s, v)$ is not available to the players, but only a perturbed observation $\widetilde{g}(s, v, \xi)$ with ξ representing some noise. Let $\mathcal{H}_n = (S \times V)^n$ denote the space of all possible sequences of moves up to time n. The general framework can be summarized as follows.

(1) The sequence $\{(s_n, v_n)\}_{n\geq 1}$ and $\{\xi_n\}_{n\geq 1}$ are random processes defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and adapted to the filtration $\{\mathcal{F}_n\}$. Moreover, the noise sequence $\{\xi_n\}$ satisfies that for some $T > 0$, each $\varepsilon > 0$, and each $(s, v) \in S \times V$,

$$\lim_{n\to\infty} \mathbb{P}\left\{\sup_{j\geq n}\max_{t\leq T}\left|\sum_{i=m(t)}^{m(jT+t)-1}\frac{1}{i}\Delta_i g\right| \geq \varepsilon\right\} = 0, \tag{III.6}$$

where $m(t) := \max\left\{n \in \mathbb{N} : \sum_{i=1}^n \frac{1}{i} \leq t\right\}$; $\Delta_i g = g(s, v) - \widetilde{g}\,(s, v; \xi_i)$.

(2) The moves of the players are independent if provided the past information h_n, i.e., $\mathbb{P}((s_{n+1}, v_{n+1}) \in ds \times dv \mid \mathcal{F}_n) = \mathbb{P}(s_{n+1} \in ds \mid \mathcal{F}_n)\mathbb{P}(v_{n+1} \in dv \mid \mathcal{F}_n)$;

(3) Player 1 only know the law $F : \mathcal{O} \times S \times V \to \mathbb{R}^m$, some bounded measurable map, of interest functions $\mathbf{X}_n$ and the next moves. The process $\mathbf{X}_n$ is defined as:

$$\mathbf{X}_{n+1} = \mathbf{X}_n + a_{n+1}F(\mathbf{X}_n, s_{n+1}, v_{n+1}, \xi_{n+1}); \tag{III.7}$$

For example, in practice, we often choose $a_n = \dfrac{1}{n}$, $\mathbf{X}_n = \dfrac{1}{n}\sum_{i=1}^n \widetilde{g}(s_i, v_i, \xi_i)$.

(4) Player 1's decisions are based on these information, i.e., $\mathbb{P}\left(s_{n+1} \in ds \mid \mathcal{F}_n\right) = Q_{\mathbf{x}_n}(ds)$, where for each $\mathbf{x} \in \mathcal{O}, Q_{\mathbf{x}}(\cdot)$ is a probability measure over S, and for each measurable set $A \subset S$, the map: $\mathbf{x} \in \mathcal{O} \to Q_{\mathbf{x}}(A) \in [0, 1]$ is measurable. The family $Q = \{Q_{\mathbf{x}} : \mathbf{x} \in \mathcal{O}\}$ is a strategy for player 1.

We first consider the case when

$$a_n = \frac{1}{n} \text{ and } \mathbf{X}_n = \frac{1}{n}\sum_{i=1}^n \widetilde{g}(s_i, v_i, \xi_i). \tag{III.8}$$

Define the map $G : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ $G(\mathbf{x}) = \overline{\text{co}}\{-\mathbf{x} + \int_S g(s, v)Q_{\Pi_{\mathcal{O}}(\mathbf{x})}(ds); v \in V\}$. Then we also get the same results as in Multistage Decision Making with Partial Observations model in Section III-A.

We now consider procedures, where, after each stage n player 1 is uninformed of his previous sequences of moves but only know the law F and player 2's moves. The new process $\mathbf{X}^*_n$ is defined as:

$$\mathbf{X}^*_{n+1} = \mathbf{X}^*_n + a_{n+1}\int_S F(\mathbf{X}^*_n, s, v_{n+1}, \xi_{n+1})Q_{\mathbf{x}^*_n}(ds). \tag{III.9}$$

Then, a procedure in law is a strategy $Q = \{Q_{\mathbf{x}^*} : \mathbf{x}^* \in \mathcal{O}\}$ as above. We will analyze a concrete case when $a_n = \dfrac{1}{n}$ and

$$\mathbf{X}^*_n = \frac{1}{n}\sum_{i=1}^n \int_S \widetilde{g}(s, v_i, \xi_i)Q_{\mathbf{x}^*_n}(ds). \tag{III.10}$$

Let $\mathbf{X}^{*n}(t)$ be the shift sequence of functions of the piecewise linear interpolation generated by the sequence $\{\mathbf{X}^*_n\}$ and the step-size sequence $\{\frac{1}{n}\}$. We also get following results about approachability.

Theorem III.5. *Suppose that $\mathbf{X}^*(\cdot)$ is the limit of a convergent subsequence of $\{\mathbf{X}^{*n}(\cdot)\}$. We get that $\mathbf{X}^*(\cdot)$ is a solution of the differential inclusion*

$$\dot{\mathbf{X}} \in G(\mathbf{X}). \tag{III.11}$$

Proof. We obtain from (III.10) that

$$\begin{aligned}(n+1)(\mathbf{X}^*_{n+1} - \mathbf{X}^*_n) &= -\mathbf{X}^*_n + \int_S g(s, v_{n+1})Q_{\mathbf{x}^*_n}(ds)\\ &+ \int_S \widetilde{g}(s, v_{n+1}, \xi_{n+1})Q_{\mathbf{x}^*_n}(ds) - \int_S g(s, v_{n+1})Q_{\mathbf{x}^*_n}(ds).\end{aligned} \tag{III.12}$$

We can check that $-\mathbf{X}^*_n + \int_S g(s, v_{n+1})Q_{\mathbf{x}^*_n}(ds) \in G(\mathbf{X}^*)$, and the remaining terms in (III.12) satisfy Assumption 1. By applying Theorem II.1, we complete the proof. □

The following theorems can be obtained as in section III-A.

Theorem III.6. *If there is a strategy Q such that E is a globally asymptotically stable set of differential inclusion (III.11), then E is approachable.*

Theorem III.7. *If the "convergence to 0" condition (III.6) is relaxed as* $|g(s,v) - \widetilde{g}(s,v,\xi)| < \eta, \forall e, b, \xi$, *w.p.1, then theorems III.5 and III.6 still hold with* G *in (III.11) being replaced by its neighbor with radius* η. *Moreover, if* E_η *is a globally asymptotically stable set of the corresponding (limit) differential inclusions (and thus, is an approachable set), then there is a (deterministic) non-decreasing function* $\phi(\cdot)$ *satisfying* $\lim_{t\to 0}\phi(t) = 0$ *such that* $\text{distance}\,(E_\eta, E) \le \phi(\eta)$.

Theorem III.8. *If there exists a strategy* Q *such that* $E = \{\mathbf{x}^*\}$ *is a unique approachable set, then under Assumption 2, the limit process of convergent subsequences of shifted and interpolated processes* $\frac{\mathbf{X}_n - \mathbf{x}^*}{\sqrt{n}}$ *converges weakly to the solution of a stochastic differential inclusion.*

Now, suppose that player 1 uses a procedure in law. Define

$$G^2(\mathbf{X},\mathbf{X}^*) = \overline{\text{co}}\{(-\mathbf{x} + \int_S g(s,v)Q_{\Pi_{\mathcal{O}}(\mathbf{x})}(ds), \int_S -\mathbf{x}^* + \int_S g(s,v)Q_{\Pi_{\mathcal{O}}(\mathbf{x}^*)}(ds)); v \in V\}.$$

Similarly, let $\mathbf{X}^n(t)$ be the shift sequence of functions of the piecewise linear interpolation generated by the sequence $\{\mathbf{X}_n\}$ and the step-size sequence $\{\frac{1}{n}\}$ and $\mathbf{X}(\cdot)$ be the limit of a convergent subsequence of $\{\mathbf{X}^n(\cdot)\}$. Applying results in Section II, we get the following theorem.

Theorem III.9. *The coupled system* ($\mathbf{X}^n, \mathbf{X}^{*n}$) *are solutions of the differential inclusion* $\left(\dot{\mathbf{X}}, \dot{\mathbf{X}}^*\right) \in G^2\left(\mathbf{X}, \mathbf{X}^*\right)$.

Thanks to the limit theorem, we then follow similarly argument in [5] to obtain the following results.

Assumption 3. G *is an upper semicontinuous correspondence from* $\mathbb{R}^m$ *to itself, with compact convex nonempty values and which satisfies the following growth condition. There exists* $c > 0$ *such that for all* $x \in \mathbb{R}^m$, $\sup_{z\in M(x)}\|z\| \le c(1 + \|x\|)$.

Assumption 4. The map F satisfies one of the two following conditions:

(i) There exists a norm $\|\cdot\|$ such that $\mathbf{x} \to \mathbf{x} + F(\mathbf{x}, s, v)$ is contracting uniformly in $h = (s,v)$. That is, $\|\mathbf{x} + F(\mathbf{x},h) - (\mathbf{y} + F(\mathbf{y},h))\| \le \rho\|\mathbf{x} - \mathbf{y}\|$ for some $\rho < 1$.

(ii) F is C^1 in $\mathbf{x}$ and there exists $\alpha > 0$ such that all eigenvalues of the symmetric matrix $\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x},h) + \frac{\partial F^T}{\partial \mathbf{x}}(\mathbf{x},h)$ are bounded by $-\alpha$, where, ^{T} stands for the transpose.

Theorem III.10. *Assume that* $\{\mathbf{X}_n, \mathbf{X}_n^*\}$ *is a bounded sequence. Under a procedure in law, the limit sets of* $\mathbf{X}^n$ *and* $\mathbf{X}^{*n}$ *coincide, and this limit set is an internally chain transitive set of the differential inclusion* (*III.11*).

IV. Concluding Remarks

We have focused on set-valued stochastic approximation in this paper. The main effort is to demonstrate the utility of the results of SA to applications. It is conceivable that such set-valued SA will make substantial impact for a wide range of applications for years to come.

References

[1] A.E. Albert and L.A. Gardner, *Stochastic Approximation and Nonlinear Regression*, MIT Press, Cambridge, MA, 1967.

[2] J.A. Bather, Stochastic approximation: A generalization of the Robbins-Monro procedure, in *Prod. 4th Prague Symposium Asymptotic Statist.*, P. Mandl and M. Hušková, eds., 13–27, 1989.

[3] M. Benaïm and M. W. Hirsch, Mixed equilibria and dynamical systems arising from fictitious play in perturbed games, *Games Econom. Behav.*, **29** (1999), 36–72.

[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, *Mathematics of Operations Research*, **31** (2006), 673–695.

[6] A. Benveniste, M. Métivier and P. Priouret, *Adaptive Algorithms and Stochastic Approximation*, Springer-Verlag, Berlin, 1990.

[7] D. Blackwell, An analog of the minmax theorem for vector payoffs, *Pacific J. Math.*, **6** (1956), 1–8.

[8] G. Brown, Iterative solution of games by fictitious play, *Activity Analysis of Production and Allocation*, T. Koopmans ed., Wiley, New York, 1951, 374–376.

[9] H.F. Chen and Y.M. Zhu, *Stochastic Approximation*, Shanghai Sci. & Tech. Publisher, Shanghai, 1996.

[10] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.

[11] M. Duflo *Random Iterative Models*, Springer-Verlag, New York, 1997.

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

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

[14] J. Kiefer and J. Wolfowitz, Stochastic estimation of the maximum of a regression function, *Ann. Math. Statist.* **23** (1952), 462–466.

[15] H.J. Kushner, An averaging method for stochastic approximations with discontinuous dynamics, constraints, and state-dependent noise, in *Recent Advances in Stochastics*, M.H. Rizri, J.S. Rustagi and D.Siegmund, eds., Academic Press, New York, 1983.

[16] H.J. Kushner and D. Clark, *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Berlin, Springer, 1978.

[17] H.J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd Ed., New York, Springer-Verlag, 2003.

[18] L. Ljung, Analysis of recursive stochastic algorithms, I*EEE Trans. Automat. Control* **22** (1977), 551–575.

[19] M. Métivier and P. Priouret, Théorèmes de convergence presque sure pour une classe dálgorithmes stochastiques à pas décroissant, *Probab. Theory Related Fields* **74** (1987), 403–428.

[20] M.B. Nevel'son, R.Z. Khasminskii, *Stochastic Approximation and Recursive Estimation*, Transl. Math. Mono., v47, AMS, Providence, 1976.

[21] N.N. Nguyen, G. Yin, Stochastic Approximation with Discontinuous Dynamics, Differential Inclusions, and Applications, *Annals of Applied Probability*, **33** (2023), 780–823.

[22] H. Robbins and S. Monro, A stochastic approximation method, *Ann. Math. Statist.* **22** (1951), 400-407.

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

[24] S. Smale, The prisoner's dilemma and dynamical systems associated to non-cooperative games, *Econometrica*, **48** (1980), 1617–1633.

[25] V. Solo and X. Kong, *Adaptive Signal Processing Algorithms*, Prentice-Hall, Englewood Clffs, NJ, 1995.

[26] S. Sorin, *A First Course on Zero-Sum Repeated Game*s, Springer, New York, 2002.

[27] J.C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation, *IEEE Trans. Automat. Control* **AC-37** (1992), 331–341.

[28] J.N. Tsitsiklis, D.P. Bertsekas, and M. Athans, Distributed asynchronous deterministic and stochastic gradient optimization algorithms, *IEEE Trans. Automat. Control* **AC-31** (1986), 803–812.

[29] Ya.Z. Tsypkin, *Adaptation and Learning in Automatic Systems*, Academic Press, New York, 1971.

[30] M.T. Wasan, *Stochastic Approximation*, Cambridge Press, London, 1969.

[31] G. Yin, V. Krishnamurthy, and C. Ion, Iterate-averaging sign algorithms for adaptive filtering with applications to blind multiuser detection, *IEEE Trans. Inform. Theory* **49** (2003), 657–671.