2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Stochastic Approximation for Nonlinear Discrete Stochastic Control: Finite-Sample Bounds for Exponentially Stable Systems

Hoang Nguyen [1] Siva Theja Maguluri [1]

Abstract— **We consider a nonlinear discrete stochastic control system, and our goal is to design a feedback control policy in order to lead the system to a prespecified state. We adopt a Stochastic Approximation (SA) viewpoint of this problem. It is known that by solving the corresponding continuous-time deterministic system, and using the resulting feedback control policy, one ensures almost sure convergence to the prespecified state in the discrete system. In this paper, we adopt such a control mechanism and provide its finite-sample convergence bounds whenever a Lyapunov function is known for the continuous system. In particular, we establish the rate $O(1/\varepsilon)$ to guarantee that the mean square error is less than ε where the Lyapunov function for the continuous system is non-smooth and gives exponential rates. Our proof relies on constructing a Lyapunov function for the discrete system based on the given Lyapunov function for the continuous system, and then appropriately smoothing the given function using the Moreau envelope. We present a numerical experiment in the selector control example to validate the established rate.**

I. Introduction

In this paper, we consider the problem of controlling a nonlinear discrete stochastic system of the form,

$$x_{k+1} = x_k + \alpha_k(F(x_k, u_k) + w_k)\,\forall\, k \in \mathbb{Z}, k \geq 0 \quad (1)$$

where $x_k \in \mathbb{R}^d$ is the state vector, $u_k \in \mathbb{R}^m$ is the control, w_k is the noise, $F(\cdot,\cdot)$ is in general a nonlinear mapping, and α_k is a sequence of step-sizes. The goal is to pick the control sequence u_k in order to ensure that the system reaches a pre-specified state x^*. We will focus on feedback control strategies of the form $u_k = \rho(x_k)$ for some mapping $\rho : \mathbb{R}^d \to \mathbb{R}^m$ to reach the state x^*. In this paper, we provide finite-time convergence bounds on the error, $\|x_k - x^*\|$ depending on the choice of the step-size sequence α_k.

Stochastic recursions of the form (1) are studied under the name of Stochastic Approximation (SA) [1], and were first introduced by Robins and Monro [2]. Asymptotic behavior of such recursions is well understood [3], [4], [1] in terms of the behavior of the corresponding continuous-time deterministic control system,

$$\dot{x} = F(x, u)\,\forall\, x \in \mathbb{R}^d \quad (2)$$

In particular, it is known that [1] under appropriate assumptions on the system, noise and choice of step-sizes, the almost-sure asymptotic behavior of the discrete-stochastic system (1) is identical to that of the continuous-deterministic system (2). Then, in order to lead the system to state x^*, one would find the optimal feedback solution $u = \rho(x(t))$ for the continuous system (2) and use the same solution for the discrete system (1). The objective of this paper is to characterize the finite-time convergence error in such an approach.

Naturally, the convergence rate of the discrete system depends on that of the continuous system. The convergence behavior as well as the rate of convergence of the continuous system is usually studied using Lyapunov arguments. In this paper, we characterize the convergence rate of the discrete system (1) based on the properties of the Lyapunov function of the continuous system (2). Suppose that there exists a feedback control policy and the corresponding Lyapunov function V for the continuous system that satisfies

$$\frac{dV}{dt} = \langle \nabla V(x), F(x) \rangle \leq -\gamma V(x)\,\forall\, x \in \mathbb{R}^d \quad (3)$$

for some $\gamma > 0$. This assumption is also known as the global exponential stability [5]. The convergence rate of the discrete system was established in the literature (in the context of optimization [6] and reinforcement learning [7]) when $V(\cdot)$ is smooth (i.e., has Lipschitz gradients). In contrast, in this paper, we consider the case when the Lyapunov function of the continuous system is non-smooth. We show that one needs $O\left(\frac{1}{\varepsilon}\right)$ samples to ensure that $\mathbb{E}[\|x_k - x^*\|^2] \leq \varepsilon$, and provide a numerical experiment with the selector control example in Section III to validate our finding.

A. Relevant literature

Finding the equilibrium point of stabilizing control problems is essentially a root-finding problem. From this perspective, the equilibrium point finding problems can be solved through the framework of SA algorithms, which were first proposed by [2]. The asymptotic convergence of SA methods was analyzed using its associated ordinary differential equations (ODE) [3], [4]. More specifically, it was shown in [1], [8] that under some conditions, the SA algorithm converges almost surely as long as the corresponding ODE is stable. In practice, it is generally preferable to have finite-sample analyses to provide performance guarantees on the output of SA algorithms after executing a finite number of operations.

In order to analyze the stability of dynamical systems, it is common to employ control Lyapunov functions. In particular, we show that the time derivative of the Lyapunov function is upper bounded by some negative constant times the Lyapunov function itself to achieve exponential stability [9]. Such conditions are found in many settings and applications in control [10], [11], [5], electrical systems [12] , robotics

[1] H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA, 30332, USA

[13] and reinforcement learning [14], [15]. Under gradient-based perspectives, the exponential stability condition is equivalent to the Polyak-Lojasiewicz condition [16] when we take the gradient flow of the iterates. Previously, [7] established finite-time analysis for nonlinear SA using an exponential dissipative assumption, which is equivalent to exponential stability condition for the ℓ_2 norm Lyapunov function. However, due to its reliance on the dissipativeness assumption, its results are limited to ℓ_2 norms and cannot be used to obtain finite-time bounds for general non-smooth nonlinear systems with a more generic potentially non-smooth Lyapunov function.

In practice, many control Lyapunov functions are non-smooth [10] since the control inputs are usually measured in discrete time rather than having continuous measurements, such as logical systems [17] or approximate discrete-time models [18]. In addition, discontinuous stabilizing feedback and non-smooth Lyapunov functions are deeply connected [19] as the non-existence of a smooth control Lyapunov function implies the absence of a continuous stabilizing feedback law [10]. To handle this, the theory of Lyapunov stability for non-smooth systems was developed by [20], where the Clarke generalized gradient [21] was used to complement the lack of gradient as we usually have in the smooth setting. The non-smooth Lyapunov analysis of equilibria is present in the differential inclusions literature [21], [22]with applications in robotics [23] and non-smooth SA [24]. A special application of non-smooth SA is the switching SA algorithm which is used in networked systems [25]. However, these prior works did not have a finite-time convergence guarantee for the non-smooth stochastic systems.

In order to deal with the non-smoothness of the system, one can attempt smoothing methods to yield a smoothened Lyapunov function. In non-smooth optimization, Moreau envelopes are commonly used [26] where the proximal is used to handle the non-smooth component of the objective [27]. In the context of SA, [28] was the first to use Moreau envelopes to obtain convergence rates for the non-smooth infinity norms, which is common for the analysis of Reinforcement Learning algorithms. However, the authors relied on the contractive property for analysis, which can be limiting as it excludes a wide range of operators without such property and many systems in control settings may not have such properties. In contrast to these prior works, our interest is to establish finite-time bounds for nonlinear SA algorithms with Lyapunov exponentially decay conditions instead of the contractive assumption for arbitrary norms.

II. System Model and Main Results

A. System model and general assumptions

Recall that we consider the problem of controlling a non-linear discrete stochastic system (1). This happens when we can only obtain the value of F via a noisy oracle $\tilde{F}$ such that for any x it will return $\tilde{F}(x,u) = F(x,u) + w$ where w is the noise (which can be dependent on the state value x and the control u). Let $\mathcal{F}_k = \{x_0, w_0, ..., x_{k-1}, w_{k-1}, x_k\}$ where $\{w_k\}$ is a martingale difference sequence with some mild conditions on its variance. We have some assumptions on the noise w_k as follows:

Assumption 1. *(Noise assumptions) The noise w_k is unbiased, that is for any $k \in \mathbb{Z}^+$:*

$$\mathbb{E}[w_k|\mathcal{F}_k] = 0 \tag{4}$$

and the noise is square-integrable. That is:

$$\mathbb{E}[\|w_k\|_2^2 |\mathcal{F}_k] \leq A + B \|x_k\|_2^2 . \tag{5}$$

In addition to these noise assumptions, we also assume that F is Lipschitz:

Assumption 2. *(Lipschitz assumption of F) There exists a positive constant C such that:*

$$\|F(x,u) - F(y,u)\|_2 \leq C \|x-y\|_2 \tag{6}$$

for any $x, y \in \mathbb{R}^d$.

In the Stochastic Gradient Descent (SGD) algorithm where $F(x) = -\nabla V(x)$, this assumption is the gradient Lipschitz assumption that is vital to ensure the stability of the algorithm. In addition, this assumption is a much more relaxed assumption than the usual contractive assumption used in [28]. Not having to rely on the contractive property will allow us to apply our results to a wider range of problems rather than just the optimal control problem with time-discounted rewards. From these assumptions, we are to analyze the discrete stochastic systems.

B. Stochastic control of non-smooth exponentially stable systems

The focus of our main results is the convergence analysis of non-smooth exponentially stable systems. If the gradient of the Lyapunov function exists everywhere, the exponential stability assumption can be written as $\dot{V}(x) = \langle \nabla V(x), \dot{x} \rangle \approx \langle \nabla V(x), F(x) \rangle \leq -\gamma V(x) \forall x \in \mathbb{R}^d$ for some positive constant γ. In absence of a smooth Lyapunov function V, we define an analogous condition using Clarke generalized gradient [21] as following: for a locally Lipschitz function V, define the generalized gradient of V at (x,t) by:

$$\partial V(x,t) = \overline{co}\{\lim \nabla V(x,t) | (x_i, t_i) \to (x,t), (x_i, t_i) \in \Omega_V\}$$

where $\overline{co}$ denotes the closed convex hull and Ω_V is the set of points where the gradient of V exists. The exponential stability assumption of the Lyapunov function can be written as:

Assumption 3. *(Exponential stability assumption of the Lyapunov function with the Clarke generalized gradient) Let $x \in \mathbb{R}^d$, we have that the following holds:*

$$\langle g_x, F(x) \rangle \leq -\gamma V(x) \, \forall \, x \in \mathbb{R}^d, g_x \in \partial V(x). \tag{7}$$

In addition, we also assume that the value of V is bounded polynomially w.r.t the distance of x to x^*, that is:

Assumption 4. *(Polynomial growth assumption of the Lyapunov function) There exists positive constants C_1, C_2 such that $\forall x \in \mathbb{R}^d$:*

$$C_1 \|x - x^*\|_2^a \le V(x) \le C_2 \|x - x^*\|_2^a. \quad (8)$$

This assumption restricts how large or how small V can be given its distance to x^* and allows us to interchangeably obtain a bound on the distance to x^* using V. Next, we have the norm of the generalized gradient of V at x is also bounded by some polynomial as well:

Assumption 5. *(Gradient growth assumption) Let $x \in \mathbb{R}^d$, for any $g_x \in \partial V(x)$, we have that:*

$$\|g_x\| \le G \|x - x^*\|_2^{a-1} \quad (9)$$

For $a = 2$, this condition implies a linear upper bound on the generalized gradient. The linear growth assumption of the gradient can be found in several works in non-smooth optimization [29] and control [9]. In the case that the gradient of the Lyapunov function is Lipschitz, it also implies linear growth of the gradient and this assumption holds automatically.

Now, we will move on to the main results. Consider $R = V^{\frac{2}{a}}$, we will show that R preserves the exponential stability condition and the norm of its gradient will be upper bounded by a linear function:

Lemma II.1. *Assume that V satisfies Assumptions 3, 4 and let $R = V^{\frac{2}{a}}$, we have that:*

$$\langle r_x, F(x)\rangle \le -\frac{2\gamma}{a} R(x) \,\forall\, x \in \mathbb{R}^d, r_x \in \partial R(x). \quad (10)$$

Furthermore, we also have that:

$$C_1^{\frac{2}{a}} \|x - x^*\|^2 \le R(x) \le C_2^{\frac{2}{a}} \|x - x^*\|^2 \;\forall\, x \in \mathbb{R}^d, \quad (11)$$

and

$$\|r_x\| \le \frac{2}{a} G C_2 \|x - x^*\| \;\forall\, x \in \mathbb{R}^d, \quad (12)$$

for $r_x \in \partial R(x)$ is the Clarke generalized gradient of R at x.

Now, we construct the Moreau envelope $M = R \square \frac{\|x\|^2}{2\mu}$ where $\square$ is the infimal convolution operator. In order to quantify how well our Moreau envelope approximates the Lyapunov function and the distance of x to x^*, we obtain some bounds on the Moreau envelope as follows:

Lemma II.2. *Let $R = V^{\frac{2}{a}}$ and $M = R \square \frac{1}{\mu} \|x\|_2^2$, we have that:*

- $C_1^{\frac{2}{a}} \|x - x^*\|^2 \le R(x) \le C_2^{\frac{2}{a}} \|x - x^*\|^2$.
- *M is convex and $\frac{1}{\mu}$-smooth.*
- $\exists 0 \le a < b : (1 + a)M(x) \le R(x) \le (1 + b)M(x)$.
- *If R defines a norm then M also defines a norm.*

Proof: Note that the first property is trivial from Assumption 4. The second and fourth properties follow from lemma 2.1 in [28]. Thus, we only need to show the third property. We have:

$$M(x) = \min_{u \in \mathbb{R}^d} \{R(u) + \frac{\|x - u\|_s^2}{2\mu}\} \le R(x), \quad (13)$$

which equality holds when $u = x$. For the LHS, note that from Assumption 4 and from $V(x)^{\frac{2}{a}}$:

$$\begin{aligned} C_1 \|x - x^*\|^a &\le V(x) \le C_2 \|x - x^*\|^a \\ \Leftrightarrow C_1^{\frac{2}{a}} \|x - x^*\|^2 &\le R(x) \le C_2^{\frac{2}{a}} \|x - x^*\|^2. \end{aligned}$$

Thus, we have:

$$\begin{aligned} M(x) &= \min_{u \in \mathbb{R}^d} \left\{ R(u) + \frac{\|x - u\|_s^2}{2\mu} \right\} \\ &\ge \min_{u \in \mathbb{R}^d} \{ C_1^{\frac{2}{a}} \|u - x^*\|_s^2 + \frac{(\|x - x^*\|_s - \|u - x^*\|_s)^2}{2\mu} \} \\ &\ge \frac{\|x - x^*\|_s^2}{\frac{1}{C_1^{\frac{2}{a}}} + 2\mu} \Rightarrow R(x) \le \left(\frac{C_2^{\frac{2}{a}}}{C_1^{\frac{2}{a}}} + 2C_2^{\frac{2}{a}} \mu \right) M(x). \quad (14) \end{aligned}$$

Lemma II.2 allows us to quantify how well M approximates R and its smoothness in terms of μ. While smaller μ gives a better approximation of R, it also scales inversely with the smoothness parameter. From here, we will show that M will also have a negative drift given a sufficiently small μ:

Lemma II.3. *There exists a constant $0 < \gamma_M < \frac{2\gamma}{a} \left(\frac{C_1}{C_2}\right)^{\frac{2}{a}}$ such that for sufficiently small $\mu > 0$:*

$$\langle \nabla M(x), F(x)\rangle \le -\gamma_M M(x) \forall x \in \mathbb{R}^d.$$

Proof: Let us denote $u = \arg\min_y \left\{ R(y) + \frac{\|x-y\|_s^2}{2\mu} \right\}$. It is well-known that $\nabla M(x) = \frac{x-u}{\mu} = g_u \in \partial R(u)$ [6]. From here, we obtain that:

$$\begin{aligned} &\langle \nabla M(x), H(x) - x\rangle = \langle g_u, F(u)\rangle + \langle g_u, F(x) - F(u)\rangle \\ &\le -\frac{2\gamma}{a} R(u) + C \|x - u\| \|\nabla M(x)\| \\ &\le -\frac{2\gamma C_1^{\frac{2}{a}} \|x - x^*\|^2}{a \left(1 + \frac{2\mu G C_1^{\frac{2}{a}-1}}{a}\right)^2} + \mu C \left(\frac{2GC_2^{\frac{2}{a}-1}}{a}\right)^2 \|u - x^*\|^2 \\ &\le \left[-\frac{2\gamma C_1^{\frac{2}{a}}}{a \left(1 + \frac{2\mu G C_1^{\frac{2}{a}-1}}{a}\right)^2} + \mu C \left(\frac{2GC_2^{\frac{2}{a}-1}}{a}\right)^2 \right] \|x - x^*\|^2 \end{aligned}$$

where the first inequality is from (10) and Assumption 2, the second inequality is from the Cauchy-Schwarz inequality and triangle inequality respectively and the last one follows from the non-expansiveness of the proximal operator. Since $0 < \gamma_M < \frac{2\gamma}{a} \left(\frac{C_1}{C_2}\right)^{\frac{2}{a}}$, we can choose μ small enough such that

$-\frac{2\gamma C_1^{\frac{2}{a}}}{a\left(1+\frac{2\mu G C_1^{\frac{2}{a}-1}}{a}\right)^2} + \mu C\left(\frac{2GC_2^{\frac{2}{a}-1}}{a}\right)^2 \leq -\gamma_M C_2^{\frac{2}{a}}$. From here, we are done since:

$$\left[-\frac{2\gamma C_1^{\frac{2}{a}}}{a\left(1+\frac{2\mu G C_1^{\frac{2}{a}-1}}{a}\right)^2} + \mu C\left(\frac{2GC_2^{\frac{2}{a}-1}}{a}\right)^2\right]\|x-x^*\|^2$$
$$\leq -\gamma_M C_2^{\frac{2}{a}}\|x-x^*\|^2 \leq -\gamma_M R(x),$$

and by the definition of Moreau envelope, note that $M(x) \leq R(x) \forall x \in \mathbb{R}^d$. Thus $\forall x \in \mathbb{R}^d$, we have $\langle \nabla M(x), F(x)\rangle \leq -\gamma_M M(x)$ for some $\gamma_M \in \left(0, \frac{2\gamma}{a}\left(\frac{C_1}{C_2}\right)^{\frac{2}{a}}\right)$.

This allows us to obtain the one-iterate bound using the smooth inequality as follows:

Proposition 1. *Suppose that V satisfies Assumptions 3, 4, and 5 and suppose that there exists $0 < \gamma_M < \frac{2\gamma}{a}\left(\frac{C_1}{C_2}\right)^{\frac{2}{a}}$ such that Lemma II.3 holds, we have:*

$$\mathbb{E}[M(x_{k+1})|\mathcal{F}_k] \leq \left(1-\frac{\alpha_k\gamma_M}{2}\right)M(x_k) + \frac{\alpha_k^2(A+2B\|x^*\|_2^2)}{\mu}. \tag{15}$$

Proof: Let $R = V^{\frac{2}{a}}$. From the Lemma II.1, we have $C_1^{\frac{2}{a}}\|x-x^*\|^2 \leq R(x) \leq C_2^{\frac{2}{a}}\|x-x^*\|^2 \forall x \in \mathbb{R}^d$ and $\langle r_x, F(x)\rangle \leq -\frac{2\gamma}{a}R(x)\forall x \in \mathbb{R}^d, r_x \in \partial R(x)$. where $\partial R(x) = \overline{co}\{\lim \nabla R(x)|x_i \to x, x_i \notin \Omega_V\}$ is the generalized gradient of R. Let $M = R\Box\frac{\|\cdot\|^2}{2\mu}$, from the $\frac{1}{\mu}$-smoothness of M, we obtain:

$$M(x_k) \leq M(x_{k-1}) + \langle \nabla M(x_{k-1}), x_k - x_{k-1}\rangle + \frac{1}{2\mu}\|x_k - x_{k-1}\|_2^2 \tag{16}$$

Taking expectations on both sides, we have:

$$\mathbb{E}[M(x_k)|\mathcal{F}_{k-1}] \leq M(x_{k-1}) + \alpha_{k-1}\langle \nabla M(x_{k-1}), F(x_{k-1})\rangle + \frac{\alpha_{k-1}^2\mathbb{E}[\|F(x_{k-1})+w_{k-1}\|^2|\mathcal{F}_{k-1}]}{2\mu}.$$

Thus, from II.3, we have that:

$$\langle \nabla M(x), F(x)\rangle \leq -\gamma_M M(x)\forall x \in \mathbb{R}^d$$

where $\gamma_M = -\frac{2\gamma C_1^{\frac{2}{a}}}{a\left(1+\frac{2\mu G C_1^{\frac{2}{a}-1}}{a}\right)^2 C_2^{\frac{2}{a}}} + \frac{C\mu}{C_1^{\frac{2}{a}}}\left(\frac{2GC_2^{\frac{2}{a}-1}}{a}\right)^2$.

Denote $E_k = \mathbb{E}[M(x_k)|\mathcal{F}_{k-1}]$, we have:

$$E_k \leq M(x_{k-1}) + \alpha_{k-1}\langle \nabla M(x_{k-1}), F(x_{k-1})\rangle + \frac{\alpha_{k-1}^2\mathbb{E}[\|F(x_{k-1})+w_{k-1}\|^2|\mathcal{F}_{k-1}]}{2\mu}.$$

This gives:

$$E_k \leq \underbrace{(1-\alpha_{k-1}\gamma_M)M(x_{k-1})}_{\text{contraction term}} + \underbrace{\frac{\alpha_{k-1}^2}{2\mu}\mathbb{E}[\|F(x_{k-1})+w_{k-1}\|_2^2|\mathcal{F}_{k-1}]}_{\text{noise term}}$$

from the fact that $F(x^*)=0$. The noise term can be further bounded as:

$$\begin{aligned}
&\frac{\alpha_{k-1}^2}{2\mu}\mathbb{E}[\|F(x_{k-1})+w_{k-1}\|_2^2|\mathcal{F}_{k-1}]\\
&\overset{(a)}{\leq} \frac{\alpha_{k-1}^2}{2\mu}\mathbb{E}[(\|F(x_{k-1})-F(x^*)\|_2+\|w_{k-1}\|_2)^2|\mathcal{F}_{k-1}]\\
&\overset{(b)}{\leq} \frac{\alpha_{k-1}^2}{2\mu}\mathbb{E}[(C\|x_{k-1}-x^*\|_2+\|w_{k-1}\|_2)^2|\mathcal{F}_{k-1}]\\
&\overset{(c)}{\leq} \frac{\alpha_{k-1}^2(C^2\|x_{k-1}-x^*\|_2^2+\mathbb{E}[\|w_{k-1}\|_2^2|\mathcal{F}_{k-1}])}{\mu}\\
&\overset{(d)}{\leq} \frac{\alpha_k^2(C^2\|x_{k-1}-x^*\|_2^2+A+B\|x_{k-1}\|_2^2)}{\mu}\\
&\overset{(e)}{\leq} \frac{\alpha_{k-1}^2(C^2+2B)\|x_{k-1}-x^*\|_2^2}{\mu} + \frac{\alpha_{k-1}^2(A+2B\|x^*\|_2^2)}{\mu}\\
&\overset{(f)}{\leq} \frac{\alpha_{k-1}^2(C^2+2B)M(x_{k-1})}{\mu\left(\frac{1}{C_1^{\frac{2}{a}}}+2\mu\right)} + \frac{\alpha_{k-1}^2(A+2B\|x^*\|_2^2)}{\mu}.
\end{aligned}$$

The first inequality (a) follows from triangle inequality, the second inequality (b) follows from Assumption 2, the third inequality (c) follows from Cauchy-Schwarz, the fourth inequality (d) follows from the Assumption 1 and the last inequality (f) follows from Lemma II.2. Thus, we have the bound on $E_k = \mathbb{E}[M(x_k)|\mathcal{F}_{k-1}]$ as:

$$E_k \leq \left(1-\alpha_{k-1}\gamma_M + \frac{\alpha_{k-1}^2(C^2+2B)}{2\mu^2+\frac{\mu}{C_1^{\frac{2}{a}}}}\right)M(x_k) + \frac{\alpha_{k-1}^2(A+2B\|x^*\|_2^2)}{\mu}.$$

Now, choose $\alpha_0 \leq \frac{\gamma_M\left(2\mu^2+\frac{\mu}{C_1^{\frac{2}{a}}}\right)}{2(C^2+2B)}$ and choose the step size sequence $\{\alpha_k\}_{k\geq 0}$ to be decreasing, we have that:

$$\mathbb{E}[M(x_{k+1})|\mathcal{F}_k] \leq \left(1-\frac{\alpha_k\gamma_M}{2}\right)M(x_k) + \frac{\alpha_k^2(A+2B\|x^*\|_2^2)}{\mu}.$$

Hence, we are done.

From the one-iterate bound, we are able to achieve the finite-time bounds of the algorithm. This can be done by expanding the one-iteration bound from the beginning to the k-th iteration and choosing a suitable step size. The bounds are summarized in the following theorem:

Theorem II.4. *Under the Assumptions 1, 2 and suppose that there exists a Lyapunov function V satisfying the Assumptions 3, 4, 5, and with the step size* $\alpha_k = \frac{\alpha}{(k+K)^\xi}$ *where* $K = \max\left\{1, \frac{\alpha(4C^2+8B)}{\gamma}\right\}$ *for* $\xi = 1$ *and* $K = \max\left\{1, \left(\frac{\alpha(4C^2+8B)}{\gamma}\right)^{\frac{1}{\xi}}, \left(\frac{2\xi}{\alpha\gamma}\right)^{\frac{1}{1-\xi}}\right\}$ *for* $\xi \in (0,1)$*, the rate of convergence for the SA problem is:*

For $\xi = 1$:

$$\mathbb{E}[\|x_k - x^*\|_2^2] \leq \begin{cases} O\left(\frac{1}{k^{\frac{\alpha\gamma_M}{2}}}\right) & \text{if } \alpha \in \left(0, \frac{2}{\gamma_M}\right) \\ O\left(\frac{\log k}{k}\right) & \text{if } \alpha = \frac{2}{\gamma_M} \\ O\left(\frac{1}{k}\right) & \text{if } \alpha \in \left(\frac{2}{\gamma_M}, \infty\right) \end{cases}$$

For $\xi \in (0,1)$:

$$\mathbb{E}[\|x_k - x^*\|_2^2] \leq C_2^{\frac{2}{a}}\left(\frac{1}{C_1^{\frac{2}{a}}} + 2\mu\right)\|x_0 - x^*\|_2^2 \times$$
$$\exp\left[-\frac{\alpha\gamma_M}{2(1-\xi)}((k+K)^{1-\xi} - K^{1-\xi})\right]$$
$$+ \frac{4\alpha}{\gamma_M(k+K)^\xi}\frac{A + 2B\|x^*\|_2^2}{\mu}\left(\frac{1}{C_1^{\frac{2}{a}}} + 2\mu\right).$$

For $\xi = 0$:

$$\mathbb{E}[\|x_k - x^*\|_2^2] \leq C_2^{\frac{2}{a}}\left(\frac{1}{C_1^{\frac{2}{a}}} + 2\mu\right)\|x_0 - x^*\|_2^2 \times$$
$$\left(1 - \frac{\alpha\gamma_M}{2}\right)^k + \frac{2(A + 2B\|x^*\|_2^2)\alpha}{\mu\gamma_M}\left(\frac{1}{C_1^{\frac{2}{a}}} + 2\mu\right).$$

Proof sketch: We expand the one-iterate bound in Proposition 1 to obtain:

$$\mathbb{E}\left[M_{\mu_{k+1}}(x_{k+1})|\mathcal{F}_k\right] \leq \underbrace{\prod_{i=1}^{k}(1 - \nu\alpha_k)}_{T_1} M_{\mu_0}(x_0)$$
$$+ \left(A^* + B^*\|x^*\|^2\right)\underbrace{\sum_{i=0}^{k}\frac{\alpha_i^2\prod_{j=i+1}^{k}(1-\nu\alpha_j)}{\mu_i}}_{T_2}.$$

The T_1 term represents how fast the RHS vanishes with each $1-\nu\alpha_k$ term is the diminishing factor at each iteration while the T_2 term represents the impact of noise in our bound. By appropriately choosing the step size α_k and with proper bounds, we obtain the finite-time bounds.

Note that since we can treat M as a smooth Lyapunov function, we can apply similar proof techniques to obtain the finite-time convergence for the smooth exponentially stable setting. From the finite-time bounds, we can easily extend the result to obtain almost sure convergence results. A detailed description of the theorem and the proof of the following corollary is in the Appendix of [30].

Corollary II.4.1. *Let* $\{x_k\}_{k\geq 0}$ *be the sequence of iterates generated by the update rule* (1)*, then when the stepsize sequence* $\{\alpha_k\}_{k\geq 0}$ *satisfies* $\sum_{k=1}^{\infty}\alpha_k = +\infty$ *and* $\sum_{k=1}^{\infty}\alpha_k^2 < +\infty$*, we have* $\lim_{k\to\infty} dist(x_k, \mathcal{X}) = 0$ *almost surely.*

III. Numerical experiments

In practical applications of control systems, oftentimes we will have discrete measurements which make our control systems non-smooth. Furthermore, external factors such as numerical precision, hardware, and environmental issues can yield noisy measurements. One such example is the selector control (Example 4.4 in [31]). Consider the system as shown in Figure 1a with the dynamic $\dot{x} = Ax + Bu$, we can write the closed-loop dynamic as:

$$\begin{aligned}\dot{x} &= Ax + B\min\{k_1^T x, k_2^T x\} \\ &= (A + Bk_1^T)x + B\min\{0, k^T x\},\end{aligned}$$

where $k = k_2 - k_1$. Let $A_1 = A + Bk_1, A_2 = A + Bk_2$ and consider the system:

$$A_1 = \begin{bmatrix}-5 & -4\\ -1 & -2\end{bmatrix}, B = \begin{bmatrix}-3\\ -21\end{bmatrix}, k = \begin{bmatrix}1\\ 0\end{bmatrix}.$$

The system is piecewise linear hence it is straightforward to see that this system satisfies Assumption 2. Note that there is no global quadratic Lyapunov function for this system, hence we have to choose a piecewise quadratic Lyapunov, which is non-smooth, in order to take into account the hybrid nature of the system:

$$V = \begin{cases} x^T P x & \text{if } k^T x \leq 0 \\ x^T(P + \eta k k^T)x & \text{otherwise}\end{cases}$$

where: $P = \begin{bmatrix}1 & 0\\ 0 & 3\end{bmatrix}, \eta = 9$. One can check that $P, P' = P + \eta k k^T$ are symmetric positive definite matrix that satisfies:

$$\begin{aligned}(A + Bk_1^T)P + P(A + Bk_1^T)^T &< 0,\\ (A + Bk_2^T)P' + P'(A + Bk_2^T)^T &< 0,\end{aligned}$$

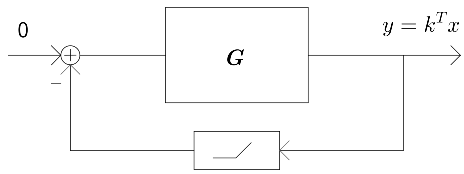
which implies that there exists $\lambda, \lambda' > 0$ such that $\dot{V} \leq -\lambda\|x\|^2 \Rightarrow \dot{V} \leq -\lambda' V$. Thus, the system satisfies Assumptions 3. Furthermore, since the Lyapunov function is piecewise quadratic and, consequently, the Clarke generalized gradient is linearly bounded. Hence, the Assumptions 4 and 5 are satisfied for $a = 2$. Applying the SA algorithm, we obtain results in Figure 1b. Note that the slope value -1.03 of the linear regressor of the log plot in 1b roughly matches the complexity $O\left(\frac{1}{k}\right)$ for $\xi = 1$.

IV. Conclusion and Future works

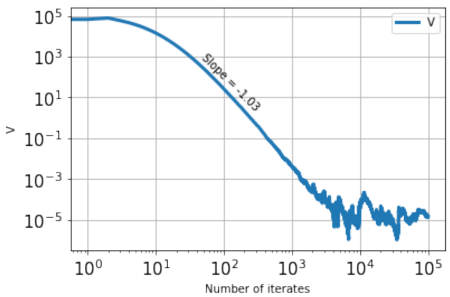
In this work, we investigated the SA framework for the non-linear discrete stochastic systems and showed that a sample complexity of $O\left(\frac{1}{\varepsilon}\right)$ is required to obtain an ε-approximation solution. A natural extension of our work is to consider more generalized stability conditions as suggested in [9] and apply the SA framework to other control settings.

V. Acknowledgements

This work was partially supported by NSF grants EPCN-2144316 and CPS-2240982.



(a) Selector control illustrations in the feedback control form



(b) The value of the Lyapunov function when $\xi = 1$

Fig. 1: An example of a non-smooth exponentially stable system with the elector control example.

References

[1] V. S. Borkar, "Stochastic approximation: A dynamical systems viewpoint," 2008.

[2] H. Robbins and S. Monro, "A stochastic approximation method," *Annals of Mathematical Statistics*, vol. 22, pp. 400–407, 1951.

[3] A. Benveniste, M. Metivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations*, 1st ed. Springer Publishing Company, Incorporated, 2012.

[4] H. Kushner and D. Clark, "Stochastic approximation methods for constrained and unconstrained systems," *SIAM Review*, vol. 22, no. 3, pp. 382–384, 1980. [Online]. Available: https://doi.org/10.1137/1022079

[5] A. D. Ames, K. Galloway, K. Sreenath, and J. W. Grizzle, "Rapidly exponentially stabilizing control lyapunov functions and hybrid zero dynamics," *IEEE Transactions on Automatic Control*, vol. 59, no. 4, pp. 876–891, 2014.

[6] A. Beck, *First-Order Methods in Optimization*. Philadelphia, PA, USA: SIAM-Society for Industrial and Applied Mathematics, 2017.

[7] Z. Chen, S. Zhang, T. T. Doan, J.-P. Clarke, and S. T. Maguluri, "Finite-sample analysis of nonlinear stochastic approximation with applications in reinforcement learning," 2019. [Online]. Available: https://arxiv.org/abs/1905.11425

[8] L. Ljung, "Analysis of recursive stochastic algorithms," *IEEE Transactions on Automatic Control*, vol. 22, no. 4, pp. 551–575, 1977.

[9] H. Khalil, *Nonlinear Control*. Pearson Education, 2014. [Online]. Available: https://books.google.com.vn/books?id=OyGvAgAAQBAJ

[10] P. Osinenko, G. Yaremenko, and G. Malaniya, "On stochastic stabilization via non-smooth control lyapunov functions," *IEEE Transactions on Automatic Control*, pp. 1–8, 2022. [Online]. Available: https://arxiv.org/abs/2205.13409

[11] Q. Nguyen and K. Sreenath, "Exponential control barrier functions for enforcing high relative-degree safety-critical constraints," in *2016 American Control Conference (ACC)*, 2016, pp. 322–328.

[12] W. Cui, J. Li, and B. Zhang, "Decentralized safe reinforcement learning for voltage control," 2021. [Online]. Available: https://arxiv.org/abs/2110.01126

[13] T. Westenbroek, F. Castaneda, A. Agrawal, S. Sastry, and K. Sreenath, "Lyapunov design for robust and efficient robotic reinforcement learning," 2022. [Online]. Available: https://arxiv.org/abs/2208.06721

[14] J. Choi, F. Castañeda, C. J. Tomlin, and K. Sreenath, "Reinforcement learning for safety-critical control under model uncertainty, using control lyapunov functions and control barrier functions," 2020. [Online]. Available: https://arxiv.org/abs/2004.07584

[15] S. Khodadadian, P. Sharma, G. Joshi, and S. T. Maguluri, "Federated reinforcement learning: Linear speedup under markovian sampling," 2022. [Online]. Available: https://arxiv.org/abs/2206.10185

[16] B. Polyak, "Gradient methods for the minimisation of functionals," *USSR Computational Mathematics and Mathematical Physics*, vol. 3, no. 4, pp. 864–878, 1963. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0041555363903823

[17] X. Yang and H. Li, "Stability analysis of probabilistic boolean networks with switching discrete probability distribution," *IEEE Transactions on Automatic Control*, pp. 1–1, 2022.

[18] P. Yu, Y. Kang, and Q. Zhang, "Sampled-data stabilization for a class of stochastic nonlinear systems with markovian switching based on the approximate discrete-time models," in *2018 Australian & New Zealand Control Conference (ANZCC)*, 2018, pp. 413–418.

[19] F. Clarke, "Lyapunov functions and discontinuous stabilizing feedback," *Annual Reviews in Control*, vol. 35, no. 1, pp. 13–33, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1367578811000022

[20] D. Shevitz and B. Paden, "Lyapunov stability theory of nonsmooth systems," *IEEE Transactions on Automatic Control*, vol. 39, no. 9, pp. 1910–1914, 1994.

[21] F. Clarke, *Optimization and Nonsmooth Analysis*. Wiley New York, 1983.

[22] J. Aubin and A. Cellina, *Differential Inclusions: Set-Valued Maps and Viability Theory*, ser. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 1984. [Online]. Available: https://books.google.it/books?id=KDqXQgAACAAJ

[23] B. E. Paden and S. S. Sastry, "A calculus for computing filippov's differential inclusion with application to the variable structure control of robot manipulators," in *1986 25th IEEE Conference on Decision and Control*, 1986, pp. 578–582.

[24] S. Majewski, B. Miasojedow, and E. Moulines, "Analysis of nonsmooth stochastic approximation: the differential inclusion approach," 2018. [Online]. Available: https://arxiv.org/abs/1805.01916

[25] G. Yin, L. Y. Wang, and T. Nguyen, "Switching stochastic approximation and applications to networked systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 9, pp. 3587–3601, 2019.

[26] J. J. Moreau, "Proximité et dualité dans un espace hilbertien," *Bulletin de la Société Mathématique de France*, vol. 93, pp. 273–299, 1965.

[27] A. Beck and M. Teboulle, "Smoothing and first order methods: A unified framework," *SIAM Journal on Optimization*, vol. 22, no. 2, pp. 557–580, 2012. [Online]. Available: https://doi.org/10.1137/100818327

[28] Z. Chen, S. T. Maguluri, S. Shakkottai, and K. Shanmugam, "Finite-sample analysis of stochastic approximation using smooth convex envelopes," 2020. [Online]. Available: https://arxiv.org/abs/2002.00874

[29] 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," 2020. [Online]. Available: https://arxiv.org/abs/2006.07904

[30] H. H. Nguyen and S. T. Maguluri, "Stochastic approximation for nonlinear discrete stochastic control: Finite-sample bounds," 2023.

[31] M. Johansson, "Piecewise linear control systems - a computational approach," in *Lecture notes in control and information sciences*, 2003.