2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Constant step-size stochastic approximation with delayed updates

Aditya Mahajan, Silviu-Iulian Niculescu, and Mathukumalli Vidyasagar

2024 IEEE 63rd Conference on Decision and Control (CDC) | 979-8-3503-1633-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/CDC56724.2024.10886824

***Abstract*— In this paper, we consider constant step-size stochastic approximation with delayed updates. For the non-delayed case, it is well known that under appropriate conditions, the discrete-time iterates of stochastic approximation track the trajectory of a continuous-time ordinary differential equation (ODE). For the delayed case, we show in this paper that, under appropriate conditions, the discrete-time iterates track the trajectory of a delay-differential equation (DDE) rather than an ODE. Thus, delayed updates lead to a qualitative change in the behavior of constant step-size stochastic approximation. We present multiple examples to illustrate the qualitative affect of delay and show that increasing the delay is generally destabilizing but, for some systems, it can be stabilizing as well.**

***Index Terms*— stochastic approximation; iterative learning algorithms; time-delay systems**

I. Introduction

Stochastic approximation refers to a family of iterative algorithms for finding the root of a function when only noisy function evaluations are available [1]. Formally, given a real function $f\colon \mathbb{R}^p \to \mathbb{R}$ and an initial value $\theta_0 \in \mathbb{R}^p$, consider the iteration

$$\theta_{n+1} = \theta_n + \alpha_n\big[f(\theta_n) + \xi_{n+1}\big], \quad n \ge 0 \tag{1}$$

where $\{\xi_{n+1}\}_{n\ge1}$ is a noise process, $\{\alpha_n\}_{n\ge1}$ is the learning rate process, and $f\colon \mathbb{R}^p \to \mathbb{R}$ is a continuous function. Stochastic approximation theory identifies conditions under which the iteration (1) converges almost surely, with high probability, or in the mean-squared sense. We refer the reader to [2]–[4] for an overview. Stochastic approximation theory is the building block of many of modern machine learning methods including system identification [5], [6], adaptive control [7], [8], reinforcement learning [9], and others.

Time heterogeneity is an important feature in a large variety of applications where transport and propagation phenomena and/or communication constraints have to be considered. In such cases, the noisy evaluation of the function is not available instantaneously but only after a delay. For example, delays appear in biological systems, financial markets, feedback communication systems, reward feedback in learning algorithms, etc. For instance, there is a lot of recent interest in multi-armed bandits (MAB) and reinforcement learning (RL) when there is a delay between taking an action and receiving a reward. Several variations of MAB with delayed reward feedback have been considered [10], [11] and follow-up literature. However, the literature on RL with delayed reward feedback is more limited [12]–[15] and a general theory characterizing the fundamental limits of convergence of RL with delayed reward feedback is still lacking.

In such settings, it is important to understand the convergence properties of the iteration

$$\theta_{n+1} = \theta_n + \alpha_n\big[f(\theta_{n-d}) + \xi_{n+1}\big], \quad n \ge d \tag{2}$$

where $d \in \mathbb{Z}_{\ge0}$ corresponds to the *processing delay*. We call this iteration *stochastic approximation with delayed updates*. For the case of $d = 0$, it reduces to the standard stochastic approximation. In this paper, we focus on the setting where the step size $\alpha_n = \alpha$ is *constant*. The standard approach to analyze such models when delay $d = 0$ is the so called *ODE approach*, which was initiated in the 1970s by Ljung [16] and extensively developed during the last 40 years, see, e.g. [3], [4], [17] and the references therein. Stochastic approximation with constant step size is useful for analysis of reinforcement learning algorithms [18] and evolutionary games and population dynamics [19], [20].

The main contribution of the paper is twofold: first, we extend the *ODE approach* for the analysis of stochastic approximation to the delay case. We will show that in the case of constant step-size, the long-term qualitative behavior of the above process is "close" to the behavior of the solutions of an appropriate delay-differential equation (DDE)[1] and thus the exponential stability of the corresponding DDE will allow concluding on the convergence properties of the stochastic approximation under consideration. The results are proposed in the case of a single delay and extend to multiple delays by following the same arguments. To the best of our knowledge, such a result represents a novelty in the literature.

Second, we show the way in which the delay, seen as a *parameter*, affects the corresponding convergence properties. More precisely, we show that increasing the delay has a dichotomic character: generally, increasing the delay has a destabilizing effect but, under some conditions, increasing the delay may have a stabilizing effect as well. Such a property, largely discussed in the analysis and control of

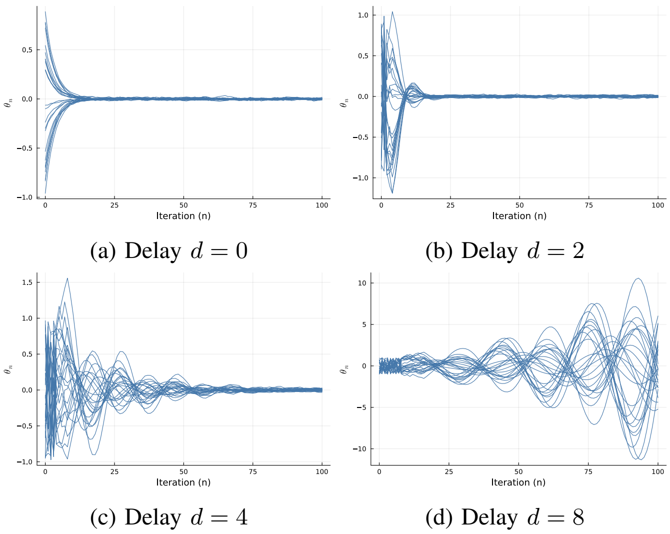
Aditya Mahajan is with the department of Electrical and Computer Engineering at McGill University, Montreal, QC H3A 0E9 Canada. Email: aditya.mahajan@mcgill.ca

Silviu-Iulian Niculescu is with University of Paris-Saclay, CNRS, CentraleSupelec, Inria Saclay, Laboratory of Signals and Systems (L2S), 91190 Gif-sur-Yvette, France. Email: silviu.niculescu@centralesupelec.fr

Mathukumalli Vidyasagar is with the department of Artificial Intelligence at the Indian Institute of Technology Hyderabad, Kandi, Telangana 502285, India. Email: m.vidyasagar@ee.iith.ac.in

The work of A. Mahajan was supported by Jean d'Alembert Fellowship from the University of Paris-Saclay and Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN-2021-0351. The work of M. Vidyasagar was supported by the Science and Engineering Research Board of India

[1] We refer the reader to [21], [22] for an overview of DDEs.



(a) Delay $d = 0$ (b) Delay $d = 2$

(c) Delay $d = 4$ (d) Delay $d = 8$

Fig. 1: Iterates for Example 1 for different values of delay.

delay systems (see, e.g. [23] and the references therein), opens interesting perspectives in stochastic approximation that need to be further explored.

In contemporaneous work, [24] have also investigated constant step-size stochastic approximation with delayed updates. They focus on finite time convergence guarantees and, therefore, their proof techniques and the nature of results are different than ours.

The remaining of the paper is organized as follows: Sec. II includes a motivating example, prerequisites and basic results in the stability analysis of DDEs. The convergence analysis of the constant step-size approximation in the case of single delays is proposed in Sec. III. Next, the proposed results are illustrated on an application involving multiple delays in Sec. IV. Finally, some concluding remarks end the paper. The notations are standard and explained when first used.

II. A MOTIVATING EXAMPLE AND BACKGROUND

To fix ideas, we present a simple example to illustrate why delay may cause a qualitative difference in the behavior of stochastic approximation.

Example 1 Consider iteration (2) for $\theta_n \in \mathbb{R}$ with $\alpha_n = \alpha = 0.1$ and

$$f(\theta) = -\kappa\theta, \quad \text{where} \quad \kappa = 2.5 \quad \text{and} \quad \xi_{n+1} \sim \mathcal{N}(0, 0.5).$$

We plot the iterates θ_n for different values of delay d. We run each simulation for $N = 100$ steps and repeat the simulation for $M = 25$ sample paths. For each sample path, the initial conditions $\theta_{1:d}$ are chosen such that θ_n, $n \in \{0, 1, \dots, d\}$, are independently sampled from $\mathrm{Unif}[-1, 1]$. The results are shown in Figure 1.

The plots show that when the delay is small, the iterates are stable and converge to zero. However, as the delay increases, the iterates become unstable and diverge. ***Thus, delay may quantitatively change the behavior of stochastic approximation.***

To understand why this is the case, recall that the standard stochastic approximation iteration (1) may be viewed as a noisy Euler discretization of the continuous-time ODE:

$$\dot{\theta}(t) = f(\theta(t)). \tag{3}$$

Therefore, under appropriate regularity conditions, the iterates of (1) converge to the equilibrium points of (3). By analogy, one may view the iterations (2) of stochastic approximation with delay as a noisy Euler discretization of the continuous-time DDE of retarded type[2]

$$\dot{\theta}(t) = f(\theta(t - \tau)), \quad \text{where } \tau = \alpha d. \tag{4}$$

Therefore, one would expect that under appropriate regularity conditions, the iterates of (2) should converge to the equilibrium points of (4).

For Example 1, Eq. (4) is a linear (scalar) DDE:

$$\dot{\theta}(t) + \kappa\theta(t - \tau) = 0, \quad \text{where } \kappa = 2.5, \tau = 0.1d. \tag{5}$$

To understand the behavior of this DDE, we review basic stability results [21], [25] for linear DDEs of the form

$$\dot{\theta}(t) = A\theta(t - \tau), \quad \text{where } \theta \in \mathbb{R}^p, A \in \mathbb{R}^{p \times p}. \tag{6}$$

The characteristic function $f : \mathbb{C} \times \mathbb{R}_+^* \mapsto \mathbb{C}$ of (6) writes as:

$$f(\lambda; \tau) = \det\left(\lambda I_p - A e^{-\lambda\tau}\right), \tag{7}$$

which is a quasipolynomial of *retarded* type[3]. Let $\{\lambda_i(A)\}_{i=1}^p$ denote the eigenvalues of A, with $|\lambda_i(A)|$ and $\angle\lambda_i(A)$ denoting their modulus and argument, respectively.

With the notations above, we have the following result (see [25] for details):

Lemma 1 *Under the assumption that A is a Hurwitz matrix, the trivial solution of the DDE (6) is exponentially stable if and only if the delay $\tau \in [0, \tau_\circ)$, where $\tau_\circ$ is given by:*

$$\tau_\circ = \min_{i \in \{1, \dots, p\}} \frac{\omega_i}{|\lambda_i(A)|}, \tag{8}$$

where $\omega_i = \min(\mathcal{A}_i^+)$ denotes the smallest positive value of the (non-empty) set $\mathcal{A}_i^+$ given by:

$$\mathcal{A}_i^+ = \mathbb{R}_+^* \cap \left\{-\frac{\pi}{2} + \angle\lambda_i(A) + 2\ell\pi : \ \ell \in \mathbb{Z}\right\},$$

for $i \in \{1, \dots p\}$.

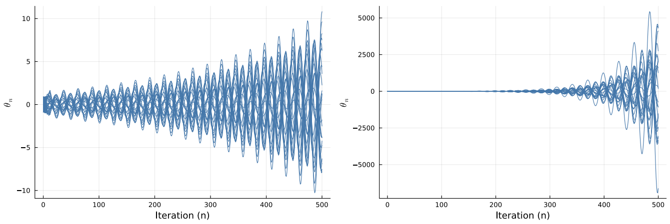
For the scalar case (5), $A = -\kappa$ (with $\kappa > 0$). So, $\lambda_1 = -\kappa = \kappa e^{i\pi}$, and $\omega_1 = \pi/2$ is the smallest positive element of the set $\mathcal{A}_1^+$. Thus, the system (4) is stable when delay $\tau = \alpha d < \tau_\circ := \pi/(2\kappa)$; thus

$$\alpha d < \frac{\pi}{2\kappa}. \tag{9}$$

The above bound is derived based on the approximate equivalence between the discrete-time iterates and the Euler discretization of the continuous time DDE, which is valid only for small values of α. It turns out that the value $\alpha = 0.1$

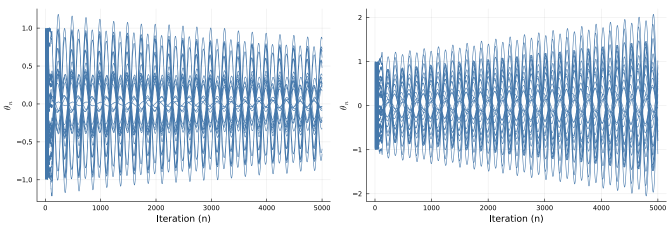
[2]For a classification of DDEs, the reader is referred to [22]; Throughout this paper, all the DDEs based-models are of retarded type.

[3]including only point spectrum, see, e.g. [25]



(a) Delay $d = 6$ (b) Delay $d = 7$

Fig. 2: Iterates for Example 1 for different values of delay close to critical delay for $\alpha = 0.1$.



(a) Delay $d = 62$ (b) Delay $d = 63$

Fig. 3: Iterates for Example 1 for different values of delay close to critical delay for $\alpha = 0.01$.

which we picked in Example 1 is not sufficiently small. For that value of α, (9) implies that the system is stable as long as $d < d_\circ := \pi/(2\alpha\kappa) = 6.283$. However, this is not the case as can be seen by the plots shown in Figure 2 for $d = \lfloor d_\circ \rfloor = 6$ and $d = \lceil d_\circ \rceil = 7$. However, if we pick a smaller value of α, say $\alpha = 0.01$, and rerun the simulation for $d = \lfloor d_\circ \rfloor = 62$ and $d = \lceil d_\circ \rceil = 63$, the results (shown in Figure 3) are consistent with the theoretical intuition.

In Sec. III, we identify general set of conditions under which the constant step-size stochastic approximation with delayed updates (2) converges.

In the example presented above, delay has a *destabilizing effect*, i.e., increasing the delay destabilizes the system. In general, delay can also have a stabilizing effect, i.e., increasing the delay stabilizes the system! In Sec. IV, we present an application which gives rise to stochastic approximation with two delay blocks and present a numerical example which highlights the stabilizing impact of delay.

III. Convergence Analysis of Constant Step-Size Stochastic Approximation with Single Delay

Consider stochastic approximation with delayed updates (2) when the step size is constant, i.e., $\alpha_n = \alpha$ for all n. The following assumptions are imposed on the model.

(A1) There exists a root θ^* of $f(\theta) = 0$.

(A2) The function f is Lipschitz continuous with a Lipschitz constant $\|f\|_L < \infty$.

(A3) The noise $\{\xi_n\}_{n\ge 1}$ is a martingale difference sequence with respect to $\mathcal{F}_n = \sigma(\theta_{1:n}, \xi_{1:n})$, i.e.,

$$\mathbb{E}[\xi_{n+1} \mid \mathcal{F}_n] = 0.$$

Moreover, $\{\xi_n\}_{n\ge 1}$ is square integrable with

$$\mathbb{E}[\|\xi_n\|^2 \mid \mathcal{F}_n] \le C_\xi(1 + \|\theta_n\|^2), \quad \text{a.s.}, n \ge 1,$$

where $C_\xi < \infty$ is a constant.

(A4) θ^* is the unique solution of $f(\theta) = 0$ and is also the global asymptotically stable equilibrium point of DDE (4).

(A5) The process $\sup_{k\in\{-d+1,\dots,0\}}\|\theta^2_{n+k}\|^2$ is uniformly integrable, i.e., there exists a constant $C_\theta < \infty$ and a Borel function $G\colon [0,\infty) \to [0,\infty)$ satisfying $G(x)/x \to \infty$ when $x \to +\infty$ such that

$$\sup_{n\ge 1} \mathbb{E}\Big[\sup_{k\in\{-d+1,\dots,0\}} \|\theta_{n+k}\|^2\Big]^{\frac{1}{2}} \le C_\theta$$

and

$$\sup_{n\ge 1} \mathbb{E}\Big[G\Big(\sup_{k\in\{-d+1,\dots,0\}} \|\theta^2_{n+k}\|^2\Big)\Big] < \infty.$$

Assumptions (A1)–(A3) are standard in the analysis of stochastic approximation [3]. Assumption (A4) is the natural analog of the corresponding stability assumption in the ODE analysis of stochastic approximation [3]. Assumption (A5) ensures that the iterates $\{\theta_n\}_{n\ge 0}$ remain stable. In principle, it is possible to relax this assumption by assuming the global asymptotic stability of an appropriately scaled DDE, following the ideas of [18], but we defer such a relaxation to future work.

For the ease of notation, we will assume that the coordinate system is chosen such that $\theta^* = 0$; i.e., under (A1), $f(0) = 0$ and under (A4), origin is global asymptotically stable solution of the DDE (4).

The high-level idea is similar to the convergence analysis of standard stochastic approximation with constant step size [3] with appropriate changes to account for the fact that the discrete-time process is tracking a DDE rather than an ODE. We start with the construction of two continuous time processes $\{\Theta(t)\}_{t\in\mathbb{R}_{\ge 0}}$ and $\{\vartheta^{(s)}(t)\}_{t\in\mathbb{R}_{\ge 0}}$ that track the discrete-time iterates $\{\theta_n\}_{n\ge 0}$ in an appropriate sense. In order to define these processes, we define the map $t(n) = n\alpha$, $n \ge 0$, which translates discrete-time to continuous-time. Let $\tau = d\alpha$ be the continuous-time processing delay.

- The process $\{\Theta(t)\}_{t\in\mathbb{R}_{\ge 0}}$ is a piecewise linear and continuous process which is defined to be equal to θ_n at time $t(n)$, $n \in \mathbb{Z}_{\ge 0}$, and is a linear interpolation between θ_n and θ_{n+1} at times in the open interval $(t(n), t(n+1))$.
- For any $t_0 \in \mathbb{R}_{\ge 0}$, the process $\{\vartheta^{(t_0)}(t)\}_{t\ge t_0}$ is the solution of the DDE (4), starting at time t_0 with the initial conditions $\vartheta^{(t_0)}(t_0+s) = \Theta(t_0+s)$, $s \in [-\tau, 0]$.

We introduce the following notations:

- For a discrete-time sequence $\{x_n\}_{n>-d}$,

$$[\![x_n]\!] := \Big[\sup_{k\in\{-d+1,\dots,0\}}\|x_{n+k}\|^2\Big]^{\frac{1}{2}}.$$

- For a continuous-time signal $\{x(t)\}_{t\ge -\tau}$,

$$[\![x_t(\cdot)]\!] := \Big[\sup_{s\in[-\tau,0]}\|x(t+s)\|^2\Big]^{\frac{1}{2}}.$$

Remark 1 By Assumption (A5), $[\![\theta_n]\!]^2$ is uniformly integrable. Therefore, there exists a positive real R such that

$$\sup_{n\geq 1} \mathbb{P}([\![\theta_n]\!] \geq R) < \alpha, \tag{10a}$$

$$\sup_{n\geq 1} \mathbb{E}\big[[\![\theta_n]\!]^2 \mathbb{1}\{[\![\theta_n]\!] \geq R\}\big] < \alpha. \tag{10b}$$

By construction, $[\![\Theta_{t(n)}(\cdot)]\!] = [\![\theta_n]\!]$. So, equation (10) also holds when $[\![\theta_n]\!]$ is replaced by $[\![\Theta_{t(n)}(\cdot)]\!]$.

Remark 2 By Assumption (A4), we can pick a $T = N\alpha$ large enough such that any solution $\vartheta(\cdot)$ of (4) starting with an initial $\vartheta(s)$, $s \in [-\tau, 0]$ with $\vartheta(0) \neq 0$, where $[\![\vartheta_0(\cdot)]\!] \leq R$ has the property that $[\![\vartheta_T(\cdot)]\!] \leq \frac{1}{2}[\![\vartheta_0(\cdot)]\!]$.

The following two lemmas are technical. Their proofs are derived by analogy to the ODE case (see, e.g., [3]) and, due to space limitations, are omitted.

Lemma 2 *For any initial time $t_0 \in \mathbb{R}_{\geq 0}$ and horizon $T \in \mathbb{R}_{\geq 0}$, we have*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} \big\|\Theta(t_0+t) - \vartheta^{(t_0)}(t_0+t)\big\|^2\Big] \in \mathcal{O}(\alpha). \tag{11}$$

Lemma 3 *For all $t \geq 0$, the following inequalities hold:*

$$\mathbb{E}\big[[\![\Theta_{t+T}(\cdot)]\!]^2 \mathbb{1}\{[\![\Theta_t(\cdot)]\!] \leq \sqrt{\alpha}\}\big]^{\frac{1}{2}} \in \mathcal{O}(\sqrt{\alpha}), \tag{12}$$

$$\mathbb{E}\big[[\![\Theta_{t+T}(\cdot)]\!]^2 \mathbb{1}\{[\![\Theta_t(\cdot)]\!] > R\}\big]^{\frac{1}{2}} \in \mathcal{O}(\sqrt{\alpha}). \tag{13}$$

With the remarks and the technical lemmas above, we can state our main result:

Theorem 1 *We have that*

$$\limsup_{n\to\infty} \mathbb{E}[\|\theta_n\|^2]^{\frac{1}{2}} \in \mathcal{O}(\sqrt{\alpha}). \tag{14}$$

Therefore, for any $\varepsilon > 0$, we have

$$\limsup_{n\to\infty} \mathbb{P}(\|\theta_n\| > \varepsilon) \in \mathcal{O}(\alpha). \tag{15}$$

PROOF Take an arbitrary n and let $t_0 = t(n)$. Then,

$$\begin{aligned}
&\mathbb{E}[[\![\theta_{n+N}]\!]^2]^{\frac{1}{2}} = \mathbb{E}[[\![\Theta_{t_0+T}(\cdot)]\!]^2]^{\frac{1}{2}} \\
&\leq \mathbb{E}[[\![\Theta_{t_0+T}(\cdot)]\!]^2 \mathbb{1}\{[\![\Theta_{t_0}(\cdot)]\!] \leq \sqrt{\alpha}\}]^{\frac{1}{2}} \\
&\quad + \mathbb{E}[[\![\Theta_{t_0+T}(\cdot)]\!]^2 \mathbb{1}\{\sqrt{\alpha} < [\![\Theta_{t_0}(\cdot)]\!] \leq R\}]^{\frac{1}{2}} \\
&\quad + \mathbb{E}[[\![\Theta_{t_0+T}(\cdot)]\!]^2 \mathbb{1}\{[\![\Theta_{t_0}(\cdot)]\!] > R\}]^{\frac{1}{2}} \\
&\overset{(a)}{\leq} \mathbb{E}[[\![\Theta_{t_0+T}(\cdot)]\!]^2 \mathbb{1}\{\sqrt{\alpha} < [\![\Theta_{t_0}(\cdot)]\!] \leq R\}]^{\frac{1}{2}} + \mathcal{O}(\sqrt{\alpha}) \\
&\overset{(b)}{\leq} \mathbb{E}[[\![\vartheta^{(t_0)}_{t_0+T}(\cdot)]\!]^2 \mathbb{1}\{\sqrt{\alpha} < [\![\Theta_{t_0}(\cdot)]\!] \leq R\}]^{\frac{1}{2}} + \mathcal{O}(\sqrt{\alpha}) \\
&\overset{(c)}{\leq} \tfrac{1}{2}\mathbb{E}[[\![\Theta_{t_0}(\cdot)]\!]^2]^{\frac{1}{2}} + \mathcal{O}(\sqrt{\alpha}) = \tfrac{1}{2}\mathbb{E}[[\![\theta_n]\!]^2]^{\frac{1}{2}} + \mathcal{O}(\sqrt{\alpha}),
\end{aligned} \tag{16}$$

where (a) follows from Lemma 3, (b) follows from Lemma 2, and (c) follows from Remark 2 and the choice of T. Let the constant corresponding to the $\mathcal{O}(\sqrt{\alpha})$ term in (16) be K. Now consider

$$\begin{aligned}
\mathbb{E}[[\![\theta_{n+2N}]\!]^2]^{\frac{1}{2}} &\leq \tfrac{1}{2}\mathbb{E}[[\![\theta_{n+N}]\!]^2]^{\frac{1}{2}} + K\sqrt{\alpha} \\
&\leq \tfrac{1}{4}\mathbb{E}[[\![\theta_n]\!]^2]^{\frac{1}{2}} + K(1+\tfrac{1}{2})\sqrt{\alpha}
\end{aligned} \tag{17}$$

where both inequalities follow from (16). Continuing this way, we can show that for any positive integer k

$$\begin{aligned}
\mathbb{E}[[\![\theta_{n+kN}]\!]^2]^{\frac{1}{2}} &\leq (\tfrac{1}{2})^k \mathbb{E}[[\![\theta_n]\!]^2]^{\frac{1}{2}} \\
&\quad + K(1 + \tfrac{1}{2} + \cdots + \tfrac{1}{2^{k-1}})\sqrt{\alpha}.
\end{aligned} \tag{18}$$

Therefore, $\limsup_{k\to\infty} \mathbb{E}[\|\theta_{n+kN}\|^2]^{\frac{1}{2}} \leq 2K\sqrt{\alpha}$. Recall that n was chosen arbitrarily. By repeating the above argument for $n+1$, $n+2$, ..., $n+N-1$, in place of n, we get

$$\limsup_{n\to\infty} \mathbb{E}[[\![\theta_n]\!]^2]^{\frac{1}{2}} \in \mathcal{O}(\sqrt{\alpha}). \tag{19}$$

Eq. (14) follows from the observation that $\|\theta_n\| \leq [\![\theta_n]\!]$, and (15) is a direct consequence of Chebyshev's inequality. ∎

IV. AN APPLICATION

Consider the problem of fine-tuning a machine learning model on a mobile device connected to a cloud server. For the purpose of this section, we abstract this problem to finding the zero of a function $f\colon \mathbb{R}^p \to \mathbb{R}$. At each time n, the mobile device and the cloud server are asked to provide an estimate of $f(\theta_n)$. The mobile device provides a "quick and dirty" answer while the cloud server provides a "slow and precise" answer. We model this situation by assuming that the mobile device (which we index by 1) returns the evaluation $f(\theta_n)$ at time $n+d_1$ with noise level $w_{1,n-d_1+1}$, while the cloud server (which we index by 2) returns the evaluation of $f(\theta_n)$ at time $n+d_2$ with noise level $w_{2,n-d_2+1}$, where $d_1 < d_2$ but the variance of $w_{1,n}$ is more than the variance of $w_{2,n}$. The computational bandwidth on the two devices is not limited. That is, if the mobile device and the cloud server have inputs θ_n and θ_{n+1} at times n and $n+1$, they will generate outputs at time $n+d_i$ and $n+d_i+1$, $i \in \{1,2\}$.

Both these evaluations are combined together as follows:

$$\begin{aligned}
\theta_{n+1} = \theta_n &+ \alpha_1\big[f(\theta_{n-d_1}) + w_{1,n-d_1+1}\big] \\
&+ \alpha_2\big[f(\theta_{n-d_2}) + w_{2,n-d_2+1}\big],
\end{aligned} \tag{20}$$

where we assume that the noise processes $\{w_{1,n}\}_{n\geq 1}$ and $\{w_{2,n}\}_{n\geq 2}$ are i.i.d. processes that are mutually independent. For the ease of notation, we take $\alpha_1 = \alpha$ and $\alpha_2 = \gamma\alpha$, and write the above iteration as

$$\theta_{n+1} = \theta_n + \alpha\big[f(\theta_{n-d_1}) + \gamma f(\theta_{n-d_2}) + \xi_{n+1}\big] \tag{21}$$

where $\xi_{n+1} = w_{1,n-d_1+1} + \gamma w_{2,n-d_2+1}$.

Following an argument similar to Sec. III, we can show that the result of Theorem 1 is also true for iteration (21), provided assumption (A4) is replaced by the following:

(A4a) θ^* is the unique solution of $f(\theta) = 0$ and is also global asymptotically stable equilibrium point of the DDE

$$\dot{\theta}(t) = f(\theta(t-\tau_1) + \gamma f(\theta(t-\tau_2)) \tag{22}$$

where $\tau_1 = \alpha d_1$ and $\tau_2 = \alpha d_2$.

Eq. (22) is a DDE with two delay blocks. We now provide two examples to illustrate the impact of the delays d_1 and d_2 and the *gain* γ on the stability of (22).

A. Scalar example with two delay blocks

Example 2 Consider iteration (21) for $\theta_n \in \mathbb{R}$ with

$$f(\theta) = -\tfrac{1}{2}\kappa\theta \quad \text{and} \quad \gamma = 1.$$

In this case, the corresponding DDE (22) is

$$\dot{\theta}(t) = -\frac{\kappa}{2}\big[\theta(t-\alpha d_1) + \theta(t-\alpha d_2)\big], \tag{23}$$

where $0 \le d_1 < d_2$. Since the delays are commensurate[4], the exponential stability of (23) is guaranteed all $\alpha \in [0, \alpha_\circ)$, with:[5]

$$\alpha_\circ = \left[d_1\kappa \operatorname{sinc}\left(\frac{\pi}{1+d_2/d_1}\right)\right]^{-1}$$

where $\operatorname{sinc}(x) = (\sin x)/x$. Furthermore, the system is unstable for all $\alpha > \alpha_\circ$.

We consider a few special cases below.

Case 1. Consider $d_1 = d_2 = d$. In this case, $\alpha_\circ = \pi/(2d\kappa)$, and the system (23) is stable iff $\alpha < \pi/(2d\kappa)$ or, equivalently, $d < d_\circ := \pi/(2\alpha\kappa)$. This recovers the result for Example 1.

Case 2. Consider now $d_1 = 1$ and $d_2 = d$. In this case, the system (23) is stable iff

$$\alpha < \frac{1}{\kappa \operatorname{sinc}\left(\dfrac{\pi}{1+d}\right)}.$$

Note that the right hand side is a decreasing function of d and converges to the limit $1/\kappa$ as $d \to \infty$ Thus, if $\alpha < \alpha_\circ := 1/\kappa$, then the system (23) is always stable for any value of d. If $\alpha > \alpha_\circ$, then there is a critical delay:

$$d_\circ := \left\lfloor \frac{\pi}{\operatorname{sinc}^{-1}(\alpha\kappa)} - 1 \right\rfloor,$$

such that (23) is stable for all $d < d_\circ$ and unstable if $d > d_\circ$.

B. Vector example with two delay blocks

Example 3 Suppose we are interested in finding the fixed point of $f(\theta) = A\theta$ based on different delayed evaluations of $f(\theta)$. Then (21) rewrites as:

$$\begin{aligned}\theta_{n+1} = \theta_n &+ \alpha\big[A\theta_{n-d_1} + w_{1,n-d_1+1}\big] \\ &+ \alpha\Gamma\big[A\theta_{n-d_2} + w_{2,n-d_2+1}\big],\end{aligned} \tag{24}$$

where $\Gamma \in \mathbb{R}^{p\times p}$.

The equivalent of assumption (A4a) in this case reads as:

(A4b) θ^* is the unique solution of $f(\theta) = 0$ and also global asymptotically stable equilibrium point of the DDE

$$\dot{\theta}(t) = A\theta(t-\tau_1) + \Gamma A\theta(t-\tau_2), \tag{25}$$

where $\tau_1 = \alpha d_1$ and $\tau_2 = \alpha d_2$.

Consider the case where $p = 2$, $d_1 = 0$, $d_2 = d > 1$, and:

$$A = \begin{bmatrix}0 & -\omega_\circ^2\\ 1 & 0\end{bmatrix} - \kappa I, \quad \Gamma = \begin{bmatrix}0 & 0\\ 0 & -\gamma\end{bmatrix},$$

[4]i.e., there is rational dependence between the delay parameters

[5]See [21], [25] for details on the analysis of such linear DDEs.

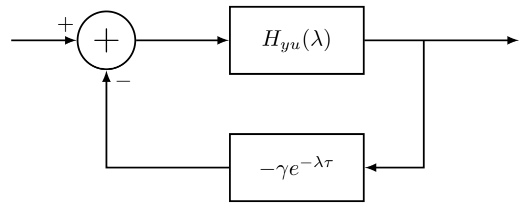


Fig. 4: Interpretation of $\Delta(\lambda)$ in (27) as the characteristic of a closed loop SISO system with time-delays.

where $\gamma, \omega_\circ \in \mathbb{R}_+^*$. Thus, the DDE (25) is equivalent to

$$\begin{aligned}\dot{\theta}(t) &= A\theta(t) + \Gamma A\theta(t-\tau)\\ &= \begin{bmatrix}-\kappa & -\omega_\circ^2\\ 1 & -\kappa\end{bmatrix}\theta(t) + \begin{bmatrix}0 & 0\\ -\gamma & \gamma\kappa\end{bmatrix}\theta(t-\tau)\end{aligned} \tag{26}$$

where $\tau = \alpha d$. The characteristic function of the DDE (26) $\Delta : \mathbb{C} \mapsto \mathbb{C}$ is given by:

$$\Delta(\lambda) := \lambda^2 + 2\lambda k + k^2 + \omega_\circ^2 - \gamma e^{-\lambda\tau}(k^2 + \lambda k + \omega_\circ^2). \tag{27}$$

It is easy to see that Δ can be interpreted as the closed-loop characteristic function of a transfer LTI SISO function

$$H_{yu}(\lambda) = \frac{\lambda k + k^2 + \omega_\circ^2}{\lambda^2 + 2\lambda k + k^2 + \omega_\circ^2}$$

controlled by a delay block $(-\gamma, \tau)$ as shown in Figure 4. In the case of low damping (κ), depending of the gain γ, the closed-loop system free of delay is exponentially stable and we will have a finite sequence of delay intervals guaranteeing the stability of the closed-loop system. In our example, we have 2 crossing frequencies $0 < \omega_- < \omega_+$ independent of the delay values: ω_+ (ω_-) corresponds to characteristic roots crossing the imaginary axis towards instability (stability) if the delay is increasing in $\mathbb{R}_+$. Furthermore, as mentioned in [25], the crossing directions at $\omega_\pm$ are independent of the delay, and these crossing roots define a partition of $\mathbb{R}_+$ in a sequence of stable delay intervals (τ_i^-, τ_{i+1}^+), and unstable delay intervals $(\tau_{i+1}^+, \tau_{i+2}^-)$, for a positive, but finite integer $i \ge 1$. Such a procedure, called *τ-decomposition method*, was introduced by Lee and Hsu [26] and it is at the origin of several approaches and procedures largely reported in the literature (see, e.g., [23] and the references therein).

In our case study, the first delay interval has the form $[0, \tau_1^+)$ where τ_1^+ represents the delay margin. Thus, to conclude, the continuous dynamical system is stable if

$$\alpha d = \tau \in [0, \tau_1^+) \cup (\tau_2^-, \tau_3^+) \cup \cdots$$

As an illustrative example, suppose $\gamma = 0.3$ and $\kappa = \gamma/4$, then the root locus is shown in Figure 5[6] with

$$\tau_1^+ = 3.4015, \quad \tau_2^- = 6.9064, \quad \tau_3^+ = 9.0049, \quad \text{etc.}$$

Note larger delays stabilize the system showing the so-called *stabilizing* effect induced by the delay, seen as a parameter. For a better understanding of such a mechanism we refer to [25] and the references therein.

[6]The authors wish thank César Fernando Méndez-Barrios (UASLP, Mexico) and Diego Torres-Garcia (UASLP, Mexico & UPSay, France) for their help in generating the graphical representation.

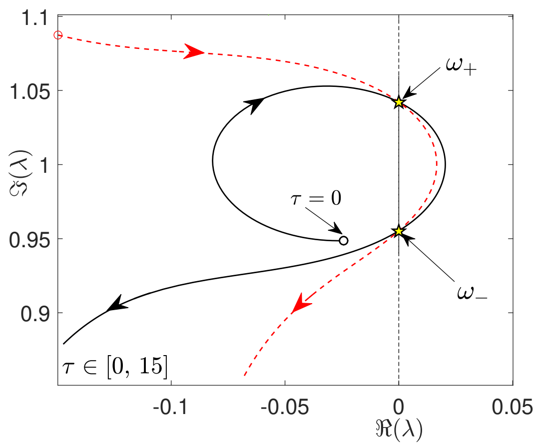


Fig. 5: Characteristic roots as a function of the delay parameter $\tau \in [0, 15]$: the black plot corresponds to the first crossing towards instability at the frequency ω_+ and back to stability at the frequency ω_- defining thus the first unstable delay interval (τ_1^+, τ_2^-); the red plot corresponds to the second crossing towards instability at the frequency ω_+ and back to stability at the frequency ω_- defining the second unstable delay interval (τ_3^+, τ_4^-), etc.

V. Conclusion

In this paper, we considered constant step-size stochastic approximation with delayed updates. Our main observation is that under appropriate conditions, the discrete-time iterates of stochastic approximation with delayed updates

$$\theta_{n+1} = \theta_n + \alpha[f(\theta_{n-d}) + \xi_{n+1}]$$

track the continuous-time trajectory of a DDE

$$\dot{\theta}(t) = f(\theta(t-\tau)), \quad \tau = \alpha d.$$

The result was derived under several technical conditions; the most critical being (A4) which asserts that the DDE should be global asymptotically stable and (A5) which asserts that the discrete-time iterates are uniformly integrable. There are several results in time-delay systems to verify the stability of DDEs. Typically, the DDE is stable for values of $\tau = \alpha d$ belonging to a union of disjoint intervals. This shows that there is a trade-off between the learning rate α and the delay d. Verifying (A5) is harder. It may be possible to generalize the scaled ODE approach of [18] to derive sufficient conditions to verify (A5), which may be an interesting future research direction.

Acknowledgments

The authors wish thank Eric Moulines (CMAP, Ecole Polytechnique, Palaiseau, France) for pointing out some of the references cited in the paper.

References

[1] H. Robbins and S. Monro, "A Stochastic Approximation Method," *The Annals of Mathematical Statistics*, vol. 22, no. 3, pp. 400 – 407, 1951.

[2] T. L. Lai, "Stochastic approximation: invited paper," *The Annals of Statistics*, vol. 31, no. 2, apr 2003.

[3] V. S. Borkar, *Stochastic approximation: a dynamical systems viewpoint*. Hindustan Book Agency, 2008.

[4] H. J. Kushner and G. G. Yin, *Stochastic Approximation Algorithms and Applications*. Springer New York, 1997.

[5] H.-F. Chen and L. Guo, *Identification and Stochastic Adaptive Control*. Birkhäuser Boston, 1991.

[6] L. Ljung, "Perspectives on system identification," *Annual Reviews in Control*, vol. 34, no. 1, pp. 1–12, 2010.

[7] Y. Tsypkin, *Adaptation and learning in automatic systems*. Academic Press: New York, 1971.

[8] P. E. Caines, *Linear stochastic systems*. SIAM, 2018.

[9] S. Meyn, *Control systems and reinforcement learning*. Cambridge University Press, 2022.

[10] N. Cesa-Bianchi, C. Gentile, Y. Mansour, and A. Minora, "Delay and cooperation in nonstochastic bandits," in *Conference on Learning Theory*, ser. Proceedings of Machine Learning Research, vol. 49. Columbia University, New York, New York, USA: PMLR, 23–26 Jun 2016, pp. 605–622.

[11] P. Joulani, A. Gyorgy, and C. Szepesvari, "Online learning under delayed feedback," in *International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 28, no. 3. Atlanta, Georgia, USA: PMLR, 17–19 Jun 2013, pp. 1453–1461.

[12] T. Lancewicki, A. Rosenberg, and Y. Mansour, "Learning adversarial markov decision processes with delayed feedback," *AAAI Conference on Artificial Intelligence*, vol. 36, no. 7, pp. 7281–7289, Jun. 2022.

[13] Y. Zhang, R. Zhang, Y. Gu, and N. Li, "Multi-agent reinforcement learning with reward delays," in *Learning for Dynamics and Control Conference*, ser. Proceedings of Machine Learning Research, N. Matni, M. Morari, and G. J. Pappas, Eds., vol. 211. PMLR, 15–16 Jun 2023, pp. 692–704.

[14] N. L. Kuang, M. Yin, M. Wang, Y.-X. Wang, and Y.-A. Ma, "Posterior sampling with delayed feedback for reinforcement learning with linear function approximation," in *Advances in Neural Information Processing Systems*, 2023.

[15] Y. Yang, H. Zhong, T. Wu, B. Liu, L. Wang, and S. S. Du, "A reduction-based framework for sequential decision making with delayed feedback," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

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

[17] M. Benaïm, "Dynamics of stochastic approximation algorithms," in *Séminaire de Probabilités XXXIII*, J. Azéma, M. Émery, M. Ledoux, and M. Yor, Eds. Berlin: Springer, 1999, pp. 1–68.

[18] V. S. Borkar and S. P. Meyn, "The o.d.e. method for convergence of stochastic approximation and reinforcement learning," *SIAM Journal on Control and Optimization*, vol. 38, no. 2, pp. 447–469, jan 2000.

[19] M. Benaïm and M. W. Hirsch, "Stochastic approximation algorithms with constant step size whose average is cooperative," *The Annals of Applied Probability*, vol. 9, no. 1, Feb. 1999.

[20] W. H. Sandholm, *Population games and evolutionary dynamics*. MIT press, 2010.

[21] R. Bellman and K. L. Cooke, *Differential-difference equations*. New York: Acadeic Press, 1963.

[22] J. K. Hale and S. M. Verduyn Lunel, *Introduction to functional-differential equations*, ser. Applied Mathematical Sciences. Springer-Verlag, New York, 1993, vol. 99.

[23] R. Sipahi, S. Niculescu, C. Abdallah, W. Michiels, and K. Gu, "Stability and stabilization of systems with time delay," *IEEE Contr. Syst. Mag.*, vol. 31, no. 1, pp. 38–65, 2011.

[24] A. Adibi, N. Dal Fabro, L. Schenato, S. Kulkarni, H. V. Poor, G. J. Pappas, H. Hassani, and A. Mitra, "Stochastic approximation with delayed updates: Finite-time rates under markovian sampling," in *Proceedings of the 27th International Conference on Artifi- cial Intelligence and Statistics (AISTATS)*, 2024.

[25] W. Michiels and S.-I. Niculescu, *Stability, Control, and Computation for Time-Delay Systems. An Eigenvalue-Based Approach (2nd ed.)*, ser. Advances in design and control. Philadelphia: SIAM, 2014.

[26] M. S. Lee and C. S. Hsu, "On the τ-decomposition method of stability analysis for retarded dynamical systems," *SIAM Journal of Control*, vol. 7, pp. 242–259, 1969.