Revisiting Step-Size Assumptions in Stochastic Approximation

Caio Kalil Lauand and Sean Meyn[*][†]

June 4, 2024

Abstract

Many machine learning and optimization algorithms are built upon the framework of stochastic approximation (SA), for which the selection of step-size (or learning rate) is essential for success. For the sake of clarity, this paper focuses on the special case $\alpha_n = \alpha_0 n^{-\rho}$ at iteration n, with $\rho \in [0, 1]$ and $\alpha_0 > 0$ design parameters. It is most common in practice to take $\rho = 0$ (constant step-size), while in more theoretically oriented papers a vanishing step-size is preferred. In particular, with $\rho \in (1/2, 1)$ it is known that on applying the averaging technique of Polyak and Ruppert, the mean-squared error (MSE) converges at the optimal rate of $O(1/n)$ and the covariance in the central limit theorem (CLT) is minimal in a precise sense.

The paper revisits step-size selection in a general Markovian setting. Under readily verifiable assumptions, the following conclusions are obtained provided $0 < \rho < 1$:

- Parameter estimates converge with probability one, and also in L_p for any $p \geq 1$.
- The MSE may converge very slowly for small ρ, of order $O(\alpha_n^2)$ even with averaging.
- For linear stochastic approximation the source of slow convergence is identified: for any $\rho \in (0, 1)$, averaging results in estimates for which the error *covariance* vanishes at the optimal rate, and moreover the CLT covariance is optimal in the sense of Polyak and Ruppert. However, necessary and sufficient conditions are obtained under which the *bias* converges to zero at rate $O(\alpha_n)$.

This is the first paper to obtain such strong conclusions while allowing for $\rho \leq 1/2$. A major conclusion is that the choice of $\rho = 0$ or even $\rho < 1/2$ is justified only in select settings—In general, bias may preclude fast convergence.

[*]This work was supported by ARO award W911NF2010055 and NSF grant CCF 2306023.

[†]Caio Kalil Lauand and Sean Meyn are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA. Emails: `caio.kalillauand@ufl.edu` and `meyn@ece.ufl.edu`

Contents

1 Introduction

Many problems of interest to the machine learning and optimization research communities hinge upon one task: root-finding in the presence of noise. That is, the goal is to estimate the vector θ^* solving $\bar{f}(\theta^*) = 0$, in which $\bar{f} \colon \mathbb{R}^d \to \mathbb{R}^d$ is defined by an expectation

$$\bar{f}(\theta) := \mathsf{E}[f(\theta, \Phi)] \tag{1}$$

where Φ is a random vector which takes values in a set X, and $f \colon \mathbb{R}^d \times \mathsf{X} \to \mathbb{R}^d$ satisfies appropriate continuity and measurability assumptions.

The standard approach is stochastic approximation (SA), defined as a d-dimensional recursion: for an initial condition $\theta_0 \in \mathbb{R}^d$,

$$\theta_{n+1} = \theta_n + \alpha_{n+1} f(\theta_n, \Phi_{n+1}) \,, \qquad n \geq 0 \tag{2}$$

in which $\{\Phi_n\}$ is a sequence of random vectors converging in distribution to Φ as $n \to \infty$ and $\{\alpha_n\}$ is known as a "step-size" sequence. A common assumption is to take the "noise" sequence $\boldsymbol{\Phi} := \{\Phi_n\}$ to be Markovian as in the setting of this paper. For ease of exposition we focus on step-size sequences of the form $\alpha_n = \alpha_0 n^{-\rho}$ in which $\alpha_0 > 0$ and $\rho \in [0, 1]$.

There is extensive theory concerning two settings: (i) vanishing step-size with $\rho \in (1/2, 1]$, and (ii) constant step-size, in which $\rho = 0$. In either case, bounds on the mean-squared error (MSE) of estimates are determined by the step-size: subject to assumptions, $\mathsf{E}[\|\theta_n - \theta^*\|^2] = O(\alpha_n)$. These assumptions are most subtle when $\rho = 1$—see the CLT for SA in [3, 6] or the MSE theory in [5, 8].

The Polyak-Ruppert (PR) averaged estimates are defined by

$$\theta_N^{\textsf{PR}} = \frac{1}{N - N_0 + 1} \sum_{k=N_0}^{N} \theta_k \,, \qquad k \geq N_0 \tag{3}$$

in which the interval $\{0, \dots, N_0\}$ is known as the *burn-in* period. Averaging was introduced in prior work to optimize the covariance in the CLT for the parameter estimates, also known as the *asymptotic covariance*. For any SA recursion this is defined by

$$\Sigma_{\theta} := \lim_{n \to \infty} n \mathsf{Cov}(\theta_n) \tag{4}$$

Letting $\Sigma^{\textsf{PR}}$ denote the asymptotic covariance for the estimates (3), under general conditions we have optimality in a sense similar to the bound of Cramèr and Rao: provided $\{\theta_n\}$ is convergent to θ^* we must have $\Sigma_\theta \geq \Sigma^{\textsf{PR}}$, where the inequality is in the matricial sense. An expression for $\Sigma^{\textsf{PR}}$ is given in (19).

A major conclusion is that the MSE convergence rate can be improved to $\mathsf{E}[\|\theta_n^{\textsf{PR}} - \theta^*\|^2] = O(1/n)$, far faster than achieved without averaging if $\rho \in [0, 1)$. *However, there is general theory only for the limited range* $\rho \in (1/2, 1)$. Recent success stories for $\rho = 0$ may be found in the literature review.

The major technical goal of this paper is to extend SA theory to the full range of possible ρ. The main results have surprising consequences to step-size selection in practice. We first recall the motivation for choosing $\rho = 0$.

Argument for constant step-size. There is intuition based on the noise-free recursion defined by $x_{n+1} = x_n + \alpha_{n+1} \bar{f}(x_n)$, $n \geq 0$, an Euler approximation of the *mean flow*,

$$\tfrac{d}{dt} \vartheta_t = \bar{f}(\vartheta_t) \tag{5}$$

Under the assumptions of this paper, the solutions to the mean flow converge to θ^* with exponential rate of convergence. We also have $x_n \to \theta^*$ geometrically fast, provided that $\rho = 0$ and $\alpha_n \equiv \alpha_0$ is chosen sufficiently small.

It is true that transients in the parameter error will show a similar decay rate for the corresponding SA recursion. This is seen by expressing (2) as a *noisy* Euler approximation of the mean flow:

$$\theta_{n+1} = \theta_n + \alpha_{n+1} [\bar{f}(\theta_n) + \Delta_{n+1}] \,, \quad n \geq 0 \tag{6}$$

with $\Delta_{n+1} = f(\theta_n, \Phi_{n+1}) - \bar{f}(\theta_n)$. In proofs of convergence as well as convergence rates for SA, an approximation of the form $\theta_n \approx \vartheta_{t_n}$ is obtained, in which $t_n = \sum_{k=1}^n \alpha_k$. For fixed step-size the convergence $\vartheta_{t_n} \to \theta^*$ holds geometrically quickly. For vanishing step-size with $0 < \rho < 1$ we have $t_n \approx \alpha_0 n^{1-\rho}/(1-\rho)$, so that the convergence rate is not geometric, but it remains *very fast*.

Of course, the approximation $\theta_n \approx \vartheta_{t_n}$ is valid only for a limited range of n—eventually noise dominates the rate of convergence.

Argument for vanishing step-size. The CLT tells us that the rate of convergence of the MSE can be no faster than $O(1/n)$, far slower than anticipated by the mean flow approximation.

A larger problem is due to *bias*. For *linear* SA, this paper obtains the the following for $\rho \in (0,1)$:

$$\begin{aligned} \theta_n^{\mathsf{PR}} &= \theta^* + \alpha_n \upbeta_n + \tfrac{1}{\sqrt{n}} \mathcal{Z}_n \\ \mathsf{E}[\|\theta_n^{\mathsf{PR}} - \theta^*\|^2] &= \alpha_n^2 \|\upbeta_n\|^2 + \tfrac{1}{n} \mathsf{E}[\|\mathcal{Z}_n\|^2] \end{aligned} \tag{7}$$

in which $\{\upbeta_n\}$ is a deterministic sequence, convergent to an identified value $\upbeta_\uptheta$, and $\{\mathcal{Z}_n\}$ is a zero-mean stochastic process satisfying $\mathsf{Cov}(\mathcal{Z}_n) \to \Sigma^{\mathsf{PR}}$ as $n \to \infty$. Consequently, if $\rho < 1/2$ and $\upbeta_\uptheta \neq 0$, then the MSE is dominated by $\alpha_n^2 \gg 1/n$.

A compatible bound is also obtained in the present paper for the general (nonlinear) SA recursion:

$$\mathsf{E}[\|\theta_n^{\mathsf{PR}} - \theta^*\|^2] \le b^{2.2} \max(\alpha_n^2, 1/n)\,, \qquad \rho \in (0,1) \tag{8}$$

This bound was previously obtained in [5] in the classical regime $\rho \in (1/2, 1)$. Under similar assumptions to the present paper they obtain:

$$\lim_{n\to\infty} n\mathsf{E}[\|\theta_n^{\mathsf{PR}} - \theta^*\|^2] = \mathrm{tr}\,(\Sigma^{\mathsf{PR}})\,, \qquad \rho \in (1/2, 1) \tag{9}$$

***Contributions*:** The main technical contributions are differentiated by the assumptions on the "noise" $\boldsymbol{\Delta} = \{\Delta_k : k \ge 1\}$. **MD**: it is a martingale difference sequence, **AD** (additive): Δ_{n+1} does not depend upon θ_n, and **MU** (multiplicative): the general setting.

In the MD setting, Δ_{n+1} may or may not depend upon θ_n. For example, the MD assumption holds when $\boldsymbol{\Phi}$ is i.i.d.. Another example of MD is Q-learning in the tabular setting [36, 28], or using split sampling [7].

The additive noise case is typical in optimization where $f(\theta_n, \Phi_{n+1}) = -\nabla L(\theta_n) + \Delta_{n+1}$, in which $L\colon \mathbb{R}^d \to \mathbb{R}$ is a loss function to be minimized, and $\boldsymbol{\Delta}$ is a zero-mean sequence (often i.i.d.).

(i) *L_p moment bounds*: for any $p \ge 1$ there exists a constant $b^{2.2}$ depending on p, ρ, and the initial condition such that for each $n \ge 1$,

$$\mathsf{E}[\|\theta_n - \theta^*\|^p] \le b^{2.2} \alpha_n^{p/2} \tag{10}$$

(ii) *Convergence and target bias*: The estimates $\{\theta_n\}$ converge to θ^* with probability one. Moreover, the rate of convergence of the average *target bias* is identified: for a constant $\bar{\Upsilon}^* \in \mathbb{R}^d$,

$$\lim_{N\to\infty} \frac{1}{\alpha_{N+1}} \mathsf{E}\Big[\frac{1}{N} \sum_{k=1}^N \bar{f}(\theta_k)\Big] = \frac{1}{1-\rho} \bar{\Upsilon}^*\,, \qquad \rho \in (0, 2/3) \tag{11}$$

in which $\bar{\Upsilon}^*$ is nonzero only in the MU setting. A representation for $\bar{\Upsilon}^*$ is given in Appendix E.

(iii) *Convergence rate*: The MSE bound (8) is established.

We are not aware of prior work establishing almost sure convergence or the MSE bound (8) *for* $\rho \in (0, 1/2]$ *even in the MD setting.*

The sharpest results are obtained for linear SA,

$$f(\theta_n, \Phi_{n+1}) = A_{n+1}\theta_n - b_{n+1} \tag{12}$$

in which the matrix A_n and vector b_n are fixed functions of Φ_n for each n. We are in the AD setting if and only if $A_n = A^*$ for each n.

(iv) The approximation (7) is established, so that PR averaging achieves the optimal asymptotic covariance for any $\rho \in (0,1)$.

***Takeaway for the practitioner*:** The Introduction leaves open who is the winner: those who advocate $\rho = 0$, or those advocating vanishing step-size with $\rho \in (1/2, 1)$. *There are no winners.*

- If your application falls in the MD or AD settings, such as typical implementations of stochastic gradient descent, then it is sensible to take a small value of ρ. The new theory in this paper demonstrates that in this case $\mathsf{E}[\|\theta_n^{\mathsf{PR}} - \theta^*\|^2] = O(1/n)$, because the bias converges rapidly: $\upbeta_n \to 0$ faster than $O(n^{-D})$ for any $D \ge 1$.

However, there is no theory in this paper that suggests $\rho = 0$ will give better performance, or adapted non-vanishing step-size such as proposed as part of ADAM [18]. Moreover, constant gain algorithms present the challenge of choosing $\alpha_0 > 0$ for stability; no such challenge presents itself for $\rho \in (0,1)$.

• If your application falls into the class MU then $\mathsf{E}[\|\theta_n^{\text{PR}} - \theta^*\|^2] = O(\|\beta_\theta\|^2 \alpha_n^2)$. The value of $\|\beta_\theta\|^2$ may be large, as illustrated in numerical experiments.

Examples of the MU setting abound in reinforcement learning. In particular, TD-learning with linear function approximation is an example of linear SA with Markovian noise.

***New approach to analysis*:** Establishing convergence rates for SA with Markovian noise often begins with the noise decomposition of Métivier and Priouret [27]:

$$\Delta_{k+1} = \mathcal{W}_{k+2} - \mathcal{T}_{k+2} + \mathcal{T}_{k+1} - \alpha_{k+1}\Upsilon_{k+2} \tag{13}$$

in which $\{\mathcal{W}_{k+2}\}$ is a martingale difference sequence and $\{\mathcal{T}_{k+2} - \mathcal{T}_{k+1}\}$ is a telescoping sequence.

The results of the present paper also rely on a noise decomposition, introduced here for the first time:

$$\Delta_{k+1} = (-\alpha_{k+1})^{m+1}\Upsilon^\bullet_{k+2} + \mathcal{W}^\bullet_{k+2} - \mathcal{T}^\bullet_{k+2} + \mathcal{R}^\bullet_{k+1} + \alpha_{k+1}[\bar{G}_{k+1}\tilde{\theta}_k - \bar{u}_{k+1}] \tag{14}$$

where $\{\mathcal{W}^\bullet_{k+2}\}$ is a martingale difference sequence, $\{\Upsilon^\bullet_{k+2}\}$ is a stochastic process with bounded L_p moments (uniform in k for each m), and the deterministic sequence of matrices $\{\bar{G}_{k+1}\}$ and vectors $\{\bar{u}_{k+1}\}$ are convergent. The sequence $\{\mathcal{T}^\bullet_{k+1} - \mathcal{R}^\bullet_{k+1}\}$ vanishes in L_p at rate α_{k+1}/k so that $\{-\mathcal{T}^\bullet_{k+2} + \mathcal{R}^\bullet_{k+1}\}$ is approximately telescoping.

This decomposition is crucial in establishing optimality of the asymptotic covariance of $\{\theta_n^{\text{PR}}\}$. First, for any given ρ we may choose m so that $(\alpha_{k+1})^{m+1} \le (\alpha_0)^m/k$, so that the first term is insignificant in covariance calculations. Second, a constant such as $\bar{u}_{k+1}$ does not change the covariance. Finally, the term $\alpha_{k+1}\bar{G}_{k+1}\tilde{\theta}_k$ may be interpreted as a vanishing perturbation of the linear dynamics. It is shown that such perturbations do not impact the asymptotic covariance—see Thm. 2.4.

1.1 Literature Survey

Asymptotic Statistics The optimal asymptotic covariance Σ^{PR} was first introduced in the 1950's for the scalar algorithm [9]. The use of averaging to achieve this lower bound appeared much later in [31, 32, 33] for general SA recursions with $\rho \in (1/2, 1)$ and $\{\Delta_k\}$ a martingale difference sequence (this is case MD in the present paper). Under these stronger assumptions on $\boldsymbol{\Delta}$, [32] provides a treatment of the regime $\rho \in (0, 1)$ for linear SA, obtaining the following conclusions for PR-averaged estimates: optimality of the CLT covariance, optimal MSE rates and almost sure convergence to θ^*.

In applications to optimization it is more common to take $\rho = 0$ [2, 37, 15, 26]. The general constant step-size algorithm with averaging is considered in [30, 12] for linear SA, where it is shown that the estimates $\{\theta_N^{\text{PR}}\}$ are convergent to θ^* and that the convergence rate is approximately optimal in a mean-square sense. Finite-n bounds are also obtained. It is assumed in [30] that $\boldsymbol{\Phi}$ is i.i.d. (independent and identically distributed), which implies the MD setting of the present paper. The paper [12] goes far further, allowing for $\boldsymbol{\Phi}$ uniformly geometrically ergodic, obtaining L_p bounds on the estimation error, and improving upon the bounds of [30] in the MD setting. We are not aware of extensions beyond the linear setting.

In broader generality, the articles [23, 14, 24] show that the PR-averaging cannot eliminate the bias induced by memory of the "noise" in the algorithm in the MU setting. The assumptions in [24] are more closely related to the present paper, in which the following bias representation was obtained for linear SA: for a constant vector $\bar{\Upsilon}^* \in \mathbb{R}^d$ and $\alpha_n \equiv \alpha_0$,

$$\lim_{N\to\infty} \mathsf{E}[\theta_N^{\text{PR}}] = \theta^* + \alpha_0[A^*]^{-1}\bar{\Upsilon}^* + O(\alpha_0^2) \tag{15}$$

See [1] for generalization to SA in which the transition kernel for $\boldsymbol{\Phi}$ is parameter dependent.

Moreover, the CLT covariance lower bound Σ^{PR} is generally not achieved in fixed step-size algorithms with averaging even for linear SA in any of the noise settings [30, 24]. In the general MU case, it admits the following approximation [24]: for a matrix $Z \in \mathbb{R}^{d\times d}$, $\lim_{N\to\infty} N\mathsf{Cov}(\tilde{\theta}_N^{\text{PR}}) = \Sigma^{\text{PR}} - \alpha_0 Z + O(\alpha_0^2)$.

In the theory of reinforcement learning the value $\rho = 1$ is often adopted, as in the original formulation of Q-learning by Watkins [35]; it was discovered in [10, 38] that this choice will result in poor convergence rate unless α_0 is chosen sufficiently large.

Finite-Time Bounds Finite-n bounds on mean square error in a Markovian setting appeared within the context of TD learning in [4]. General linear SA recursions are studied in [34], followed by [8] which focused in nonlinear recursions.

The work [8] is most closely related to the present work because they also allow for vanishing stepsize of the form $\alpha_n = \alpha_0 n^{-\rho}$. The standard SA recursion (2) is considered in [8], of the form

$$f(\theta_n, \Phi_{n+1}) = F(\theta_n, X_{n+1}) + \mathcal{W}^0_{k+1} \tag{16}$$

where $\{X_n\}$ is a Markov chain on a state space X with transition kernel P, and $\{\mathcal{W}^0_n\}$ a martingale difference sequence satisfying the standard assumptions of the SA literature [6]. It is assumed that $\{X_n\}$ is *uniformly ergodic* [29]: there is a unique invariant measure π and fixed constants $R < \infty$ and $\varrho < 1$ such that $\|P^n(x, \cdot\,) - \pi(\,\cdot\,)\|_{\text{TV}} \le R\varrho^n$.

In addition, it is assumed that $V(\theta) = \frac{1}{2}\|\theta - \theta^*\|^2$ serves as a Lyapunov function for the mean flow, in the sense that the following holds, for some $c_0 > 0$ and all θ:

$$\nabla V(\theta) \cdot \bar{f}(\theta) \le -c_0 V(\theta) \tag{17a}$$

Under these assumptions, the mean square error bound is of the form

$$\mathsf{E}[\|\tilde{\theta}_n\|^2] \le K[\log\big(n/\alpha_0\big) + 1]\,\alpha_n + \varepsilon_n\,, \qquad K = L\frac{1}{c_0}\max\Big\{1, \frac{\log(R/\varrho)}{\log(1/\varrho)}\Big\} \tag{17b}$$

in which $\{\varepsilon_n\}$ vanishes *quickly*, similar to the bound on the right hand side of (20a). The constant L takes the form $L = 520(L_{\bar{f}} + \sigma^2_{\mathcal{W}^0})^2\alpha_0(\|\theta^*\| + 1)^2$, where $L_{\bar{f}}$ is the Lipschitz constant for $\bar{f}$ and $\sigma^2_{\mathcal{W}^0}$ is a constant depending upon the variance of $\mathcal{W}^0_{k+1}$.

The present paper complements [8], in that the slow convergence for $\rho < 1/2$ is explained by the large bias in this regime when there is multiplicative noise—see (20b). Prior to the present paper, it might have been expected that (17b) could be improved using PR-averaging. The bias formula (20b) also tells us that the bound (17b) is loose due to the $\log(n)$ coefficient, but this may be a necessary price for such an elegant upper bound.

Organization: The paper is organized into three additional sections. Section 2 introduces the assumptions that are imposed throughout the paper, followed by contributions (i)–(iv). Section 3 illustrates the theory in Section 2 through a numerical experiment. Conclusions and directions for future research are included in Section 4. Technical proofs of the main results in Section 2 are contained on the Appendix.

2 Main results

2.1 Assumptions

It is assumed that $\boldsymbol{\Phi} := \{\Phi_n : n \ge 0\}$ is an aperiodic, ψ-irreducible Markov chain on a general state space X and with transition kernel P. See [29] for precise definitions of aperiodicity and ψ-ireducibility. It admits an unique invariant measure $\uppi$, so that $\bar{f}(\theta) = \mathsf{E}_\uppi[f(\theta, \Phi_k)]$, where the subscript $\uppi$ denotes the expectation is taken in steady state: $\Phi_k \sim \uppi$ for each k.

The joint parameter-disturbance process is expressed as $\boldsymbol{\Psi} := \{\Psi_n = (\theta_n, \Phi_{n+1}) : n \ge 0\}$. Any functions $g : \mathsf{X} \to \mathbb{R}^d$, $h : \mathbb{R}^d \times \mathsf{X} \to \mathbb{R}^d$ are assumed to be measurable with respect to the Borel sigma-algebra $\mathcal{B}(\mathsf{X})$ and $\mathcal{B}(\mathbb{R}^d \times \mathsf{X})$, respectively.

The following additional assumptions are imposed throughout the paper:

(A1) The SA recursion (2) is considered with $\{\alpha_n\}$ of the form $\alpha_n = \alpha_0 n^{-\rho}$ with $\rho \in (0, 1)$ and $\alpha_0 > 0$.

(A2) There exists a function $L\colon \mathsf{X} \to \mathbb{R}$ satisfying, for all $x \in \mathsf{X}$ and $\theta, \theta' \in \mathbb{R}^d$,

$$\|f(\theta, x) - f(\theta', x)\| \le L(x)\|\theta - \theta'\|$$
$$\|f(0, x)\| \le L(x)$$

(A3) The mean flow ODE (5) is globally asymptotically stable with unique equilibrium $\theta^* \in \mathbb{R}^d$ and the scaled vector field $\bar{f}_\infty(\theta) := \lim_{c\to\infty} \frac{1}{c} f(c\theta)$ exists for each $\theta \in \mathbb{R}^d$. Moreover, the ODE@∞ [7],

$$\tfrac{d}{dt}\vartheta^\infty_t = \bar{f}_\infty(\vartheta^\infty_t)$$

is globally asymptotically stable.

(A4) The Markov chain $\boldsymbol{\Phi}$ satisfies (DV3):

For functions $V\colon \mathsf{X}\to\mathbb{R}_+$, $W\colon \mathsf{X}\to[1,\infty)$, a small set C, $b>0$ and all $x\in\mathsf{X}$,

$$\left.\begin{array}{l} \mathsf{E}\big[\exp\big(V(\Phi_{k+1})\big)\mid \Phi_k=x\big]\le \exp\big(V(x)-W(x)+b\mathbf{1}_C(x)\big) \end{array}\right\}\quad \textbf{(DV3)}$$

In addition, for each $r>0$,

$$S_W(r):=\{x: W(x)\le r\}\ \text{ is either small or empty,}\quad \sup\{V(x): x\in S_W(r)\}<\infty\,,$$

$$\lim_{n\to\infty}\Big(\sup\Big\{\frac{L(x)}{W(x)}: W(x)\ge n\Big\}\Big)=0 \tag{18}$$

(A5) $\bar f:\mathbb{R}^d\to\mathbb{R}^d$ is continuously differentiable in θ, and the Jacobian matrix $\bar A=\partial\bar f$ is uniformly bounded and uniformly Lipschitz continuous with Lipschitz constant L_A. Moreover, $A^*:=\bar A(\theta^*)$ is Hurwitz.

The bound (DV3) in (A4) implies geometric ergodicity of $\boldsymbol{\Phi}$ [21]. Moreover, assumptions (A2), (A3), and (A5) imply exponential asymptotic stability of (5)[23].

Proposition 2.1. *Under (A2), (A3), and (A5), the ODE* (5) *is exponentially asymptotically stable: for positive constants $b^{2.1}$ and $\varrho^{2.1}$, and any initial condition $\vartheta_0\in\mathbb{R}^d$, $\|\vartheta_t-\theta^*\|\le b^{2.1}\|\vartheta_0-\theta^*\|\exp(-\varrho^{2.1}t)$, $t>0$.* □

2.2 General Stochastic Approximation

Moment bounds of the form (10) for general nonlinear SA recursions (2) with Markovian noise are established in [5] for $p=4$ under slightly weaker assumptions as in the present paper. In particular, the authors consider a relaxation of the condition (18) in (A4) (see Section 2.2 for more details). However, this prior work concerns the $\rho\in(1/2,1)$ regime for vanishing step-size algorithms. An extension to constant step-size is presented in [24] where L_4 moment bounds on estimation error are also established.

Similar arguments as the ones used in these prior works can be used to establish moment bounds in the setting of this paper. Under (A4), we obtain bounds of the form (10) for arbitrary p, which in turn imply contributions (ii) and (iii).

Theorem 2.2. *Suppose (A1)–(A4) hold. Then, for each $\rho\in(0,1)$,*

(i) *There exists $b^{2.2}<\infty$ depending upon Ψ_0 and ρ such that* (10) *holds.*

(ii) *The sequence of estimates converges to θ^* for each initial condition $\theta_0\in\mathbb{R}^d$ with probability one: $\theta_n\stackrel{\text{a.s.}}{\to}\theta^*$.*

If in addition (A5) holds,

(iii) *There exists $b^{2.2}$ depending upon Ψ_0 and ρ such that* (8) *holds.*

(iv) *The representation for the average target bias in* (11) *holds. Moreover, the right hand side of* (11) *is zero outside of the MU setting.*

The proofs of part (i) and (ii) of Thm. 2.2 are given at the end of Appendix B, while parts (iii) and (iv) are proven in Appendix C.

Thm. 2.2 (iv) hints at a deeper connection between performance of slowly vanishing and constant step-size algorithms. See Thm. 2.4 for a bias representation entirely analogous to (15) for linear SA, which is obtained as a consequence of (11).

Recall from Section 1.1 that L_p moment bounds and almost sure convergence were never established in prior work for $\rho\in(0,1/2]$, in any of the noise settings. Convergence has been established in only one special case: linear SA with PR-averaging in the MD setting [32].

With Thm. 2.2 in hand, it is straightforward to relax the linearity assumption, and convergence does not require averaging:

Corollary 2.3. *Suppose the assumptions of Thm. 2.2 hold. Suppose moreover that $\{\Delta_{n+1}\}$ is a martingale difference sequence and there is a constant $\overline{\sigma}_\Delta>0$ such that*

$$\mathsf{E}[\|\Delta_{n+1}\|^2|\mathcal{F}_n]\le\overline{\sigma}_\Delta^2(1+\|\theta_n\|^2)\,,n\ge0$$

where $\mathcal{F}_n:=\sigma(\theta_0,\Phi_1,\cdots,\Phi_n)$. Then, for each $\rho\in(0,1)$, parts (i)–(iv) of Thm. 2.2 hold.

The proof of Corollary 2.3 is postponed to the end of Appendix C.

2.3 Linear Stochastic Approximation

This section concerns asymptotic statistics for linear SA recursions of the form (12). In the general MU setting, it is assumed that $A_{n+1} = \mathsf{E}_\pi[A(\Phi_{n+1})] = A^*$ and $b_{n+1} = \mathsf{E}_\pi[b(\Phi_{n+1})] = \bar{b}$. Recall that $A_n = A^*$ for each n when the noise is AD. The mean vector field takes the form $\bar{f}(\theta) = A^*(\theta - \theta^*)$ with $\theta^* = [A^*]^{-1}\bar{b}$.

The optimal asymptotic covariance Σ^{PR} is defined as follows:

$$\Sigma^{\mathsf{PR}} := G\Sigma_{\mathcal{W}^*}G^{\intercal} \tag{19}$$

where $G = [A^*]^{-1}$ and $\Sigma_{\mathcal{W}^*}$ is the asymptotic covariance matrix of the martingale difference sequence $\{\mathcal{W}_k^*\}$, depending only on the "noise" $\mathbf{\Phi}$. See (40) in Appendix C for a precise definition of $\mathcal{W}^*$ in terms of algorithms primitives.

Theorem 2.4. *Suppose (A1)–(A5) hold for SA recursions of the form* (12)*. Then, for each* $\rho \in (0,1)$*,*

(i) *If the noise is AD, there exist* $b^{2.4}$*,* $\varrho^{2.4}$ *and* $n_b > 0$ *sufficiently large such that*

$$\|\mathsf{E}[\tilde{\theta}_n]\| \le b^{2.4}\exp(-\varrho^{2.4}(\tau_n^b - \tau_{n_b}^b)) \tag{20a}$$

in which $\tau_n^b = \alpha_0(1 + (1-\rho)^{-1}[n^{1-\rho} - 1])$.

(ii) *If the noise is MU and* $\rho \in (0, 2/3)$*, there exists a vector* $\upbeta_\uptheta \in \mathbb{R}^d$ *such that*

$$\lim_{n\to\infty} \frac{1}{\alpha_{n+1}}\mathsf{E}[\tilde{\theta}_n^{\mathsf{PR}}] = \upbeta_\uptheta\,, \qquad \tilde{\theta}^{\mathsf{PR}} := \theta^{\mathsf{PR}} - \theta^* \tag{20b}$$

(iii) *For any of the three noise settings (including MU),* $\lim_{n\to\infty} n\mathsf{Cov}(\theta_n^{\mathsf{PR}}) = \Sigma^{\mathsf{PR}}$

The asymptotic bias in (20b) may be expressed $\upbeta_\uptheta = (1-\rho)^{-1}[A^*]^{-1}\bar{\Upsilon}^*$, in which a representation for $\bar{\Upsilon}^*$ is given in Appendix E.

The limits in eqs. (20a) and (20b) come with error bounds:

Theorem 2.5. *Under the assumptions of Thm. 2.4 we have the following representations: for a constant* $b^{2.5}$ *depending upon* Ψ_0 *and* ρ*,*

(i)

$$\mathsf{E}[\tilde{\theta}_n^{\mathsf{PR}}] = \alpha_{n+1}\upbeta_\uptheta + \varepsilon_n^\upbeta\,, \qquad \|\varepsilon_n^\upbeta\| \le \begin{cases} b^{2.5}n^{-3\rho/2} & \rho \le 1/2 \\ b^{2.5}n^{\rho/2-1} & \rho > 1/2 \end{cases} \tag{21a}$$

(ii) *Denoting* $\sigma^{\mathsf{PR}} := \sqrt{\mathrm{tr}\,(\Sigma^{\mathsf{PR}})}$*,*

$$\|\tilde{\theta}_n^{\mathsf{PR}}\|_2 \le \alpha_{n+1}\|\upbeta_\uptheta\| + \tfrac{1}{\sqrt{n}}\sigma^{\mathsf{PR}} + \varepsilon_n^\upsigma\,, \qquad |\varepsilon_n^\upsigma| \le \begin{cases} b^{2.5}n^{-3\rho/2} & \rho \le 3/7 \\ b^{2.5}n^{-(3-\rho)/4} & \rho > 3/7 \end{cases} \tag{21b}$$

Consequently, if $\|\upbeta_\uptheta\|$ *is non-zero, then the MSE* $\|\tilde{\theta}_n^{\mathsf{PR}}\|_2^2$ *converges to zero at the optimal rate* $1/n$ *if and only if* $\rho > 1/2$*.*

The proof of parts (i) and (ii) of Thm. 2.4 are given at the end of Appendix E.1. A proof of (iii) for the AD setting can be found at the start of Appendix E.2, while the proof for the MU setting is given at the end of Appendix E.2. Thm. 2.5 is obtained by identifying the dominant error terms in Thm. 2.4.

3 Numerical Experiments

The numerical experiments contained in this section aim to illustrate the general theory, focusing on a simple linear model for which the asymptotic bias and covariance are easily computed.

An instance of the recursion (12) is considered in which $\{\Phi_k\}$ evolves on $\mathsf{X} = \{0,1\}$, with transition matrix $P = \begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix}$ where $a \in (0,1)$. The Markov chain is reversible, with uniform invariant distribution

π. Moreover, it satisfies the assumptions of [8]: it is uniformly ergodic and its rate of convergence can be identified: $\|P^n(x,\cdot) - \pi(\cdot)\|_{\text{TV}} \le R\varrho^n$ with $R = 1/2$ and $\varrho = 2a - 1$.

For a pair of matrices and a pair of vectors $A^0, A^1 \in \mathbb{R}^{d\times d}$ and $b^0, b^1 \in \mathbb{R}^d$, we have (12) with

$$A_{k+1} = \Phi_k A^1 + (1-\Phi_k)A^0 \qquad b_{k+1} = \Phi_k b^1 + (1-\Phi_k)b^0 \tag{22a}$$

Since π is uniform, the mean flow vector field is

$$\bar f(\theta) = A^*\theta - \bar b \quad \text{with } A^* = \tfrac12(A^0 + A^1)\,, \quad \bar b = \tfrac12(b^0 + b^1) \tag{22b}$$

Closed-form expressions for the terms in Thm. 2.4 are summarized in the following.

Lemma 3.1. *For $0 < a < 1$ the asymptotic statistics of the linear model are computable, with*

(i) *Asymptotic bias* (20b)*:*

$$\upbeta_\uptheta = \frac{1}{1-\rho}[A^*]^{-1}\bar\Upsilon^* \quad \textit{with} \quad \bar\Upsilon^* = \frac{(2a-1)}{4(1-a)}(A^1 - A^0)(A^0\theta^* - b^0) \tag{23}$$

(ii) *Optimal asymptotic covariance* (19)*:* $\Sigma^{\text{PR}} = \dfrac{a}{1-a}(A^0\theta^*)(A^0\theta^*)^{\intercal}$.

The proof of Lemma 3.1 is postponed to Appendix F

Setup: In the experiments surveyed here, the linear SA recursion (12) using (22a) was simulated with $a \in \{0.3, 0.5, 0.7\}$ and

$$\begin{aligned} A^0 &= 2\begin{bmatrix} -2 & 0 \\ 1 & -2 \end{bmatrix}, \quad b^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ A^1 &= 2\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, \quad b^1 = -2\begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned} \tag{24}$$

so that $\bar f(\theta) = A^*\theta - \frac12 b^1$ with $A^* = -I$ and $\theta^* = -\frac12 b^1$.

Several choices of step-size were considered, in which $\alpha_n = \alpha_0 n^{-\rho}$ with $\alpha_0 = 0.5$ and $\rho \in \{0.15, 0.3, 0.45, 0.6, 0.75, 0.9\}$. For each value of ρ, $M = 300$ independent experiments were carried out for a time horizon of $N = 3\times 10^5$ with initial conditions $\{\theta_0^i : 1 \le i \le M\}$ sampled independently from $N(\theta^*, 25I)$.

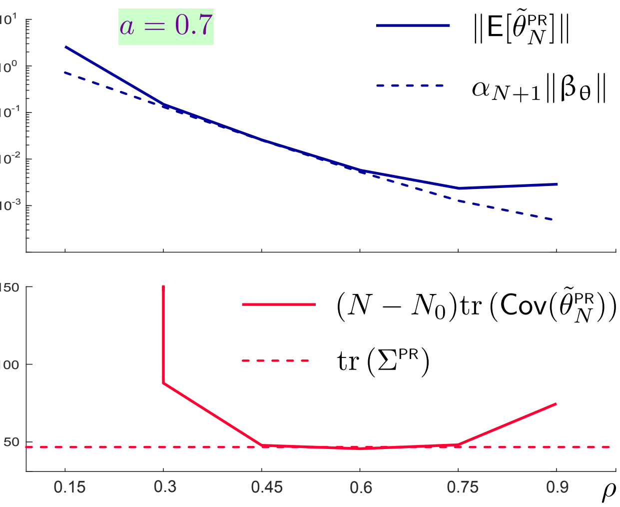

Figure 1: Empirical and theoretical mean and variance.

After obtaining the sequence of estimates $\{\theta_n^i\}$ for the i^{th} experiment, PR-averaging was applied with $N_0 = 2\times 10^3$ to compute $\{\tilde\theta_n^{\text{PR}^i} : n > N_0\}$.

Results: The empirical mean and variance were calculated for the final PR-averaged estimates $\theta_N^{\text{PR}^i}$ across all independent runs and are plotted as functions of ρ in Fig. 1 for the special case $a = 0.7$. Also plotted in this figure are their associated theoretical optimal values obtained from Thm. 2.4. Results for the remaining choices of a are displayed in Figs. 3 and 4 in Appendix F. We see that the empirical mean and covariance are near optimal for $\rho \in \{0.45, 0.6, 0.75\}$.

Small-ρ regime. Problems are observed for small ρ as expected in view of (21b) since $\upbeta_\uptheta \neq 0$ (see (23)). Moreover, for small $\rho > 0$ there is a need for a longer run time because the step-size vanishes so slowly. For example, when $\rho = 0.15$ we have that $\alpha_N \approx 0.09$ and if the same experiments were carried out with $\alpha_n \equiv \alpha_N$ for all n, one could not expect optimality of the asymptotic covariance based upon theory in [24, 30].

Large-ρ regime. Also observed in Fig. 1 is poor solidarity with theory for $\rho \sim 1$. This is predicted by Thm. 2.5 since the error terms ε_n^{β}, ε_n^{σ} converge to zero at rate $n^{-(3-\rho)/4} \approx n^{-1/2}$ when $\rho \approx 1$.

Performance without averaging. Shown on the left hand side of Fig. 2 is a plot of the approximation for the MSE $\{\|\tilde{\theta}_n\|_2^2\}$, along with the finite-$n$ bounds (17b). See Appendix F for details on how each of the terms in (17b) were identified for this example. The bound is very loose even though the maximum in (17b) defining K is equal to unity. However, remember that the bound of [8] is universal, over all nonlinear SA algorithms with common values of K.

Averaging. The plots on the right hand side of Fig. 2 show an approximation for the MSE $\{\|\tilde{\theta}_n^{\text{PR}}\|_2^2\}$, along with the approximate MSE obtained from Thm. 2.5 (obtained by dropping the error term ε_n^{σ} in (21b)). The approximation is surprisingly tight over the entire run.

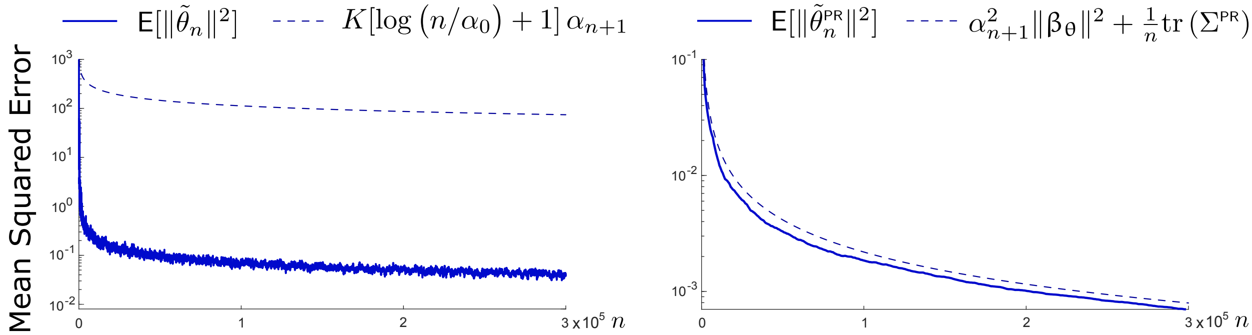

Figure 2: Comparison between empirical and theoretical mean squared error with and without PR averaging.

4 Conclusions

The condition $\rho > 1/2$ is imposed in classical treatments of stochastic approximation because of the standard requirements,

$$\sum_{n=1}^{\infty} \alpha_n = \infty\,, \qquad \sum_{n=1}^{\infty} \alpha_n^2 < \infty\,,$$

in which the first condition is essential for sufficient exploration. In the standard ODE approximation approach to convergence, the second condition is essential in analysis to bound the impact of noise in the SA recursion. This paper establishes convergence for the entire range of $\rho \in (0,1)$ based on an entirely new proof technique, in which the first step is establishing convergence rates in L_p for any p. This is also the first paper to quantify the impact of *bias* when $\rho < 1$.

It is worth repeating that the implications for the practitioner depend on the setting:

◦ In the MD or AD settings, any value of $\rho \in (0,1)$ will achieve the optimal rate of convergence of the MSE (that is, $O(1/n)$).

◦ If the noise is multiplicative (MU setting), then $\rho > 1/2$ is required to achieve the optimal rate of convergence.

There are several paths for additional research:

- It is known that $\lim_{n\to\infty} n\mathsf{E}[\|\theta_n^{\text{PR}} - \theta^*\|^2] = \mathrm{tr}\,(\Sigma^{\text{PR}})$ for $\rho \in (1/2,1)$ for nonlinear SA, even if $\bar{\Upsilon}^* \neq 0$ [5]. Can this be extended to $\rho \in (0,1/2]$ when $\bar{\Upsilon}^* = 0$? Can a version of (7) be established outside of the linear setting when $\bar{\Upsilon}^* \neq 0$?

- In applications to reinforcement learning and gradient-free optimization, the "noise" $\mathbf{\Phi}$ is designed by the user for the purpose of "exploration". Recent success stories show that it is possible to select $\mathbf{\Phi}$ so that $\mathrm{tr}\,(\Sigma^{\text{PR}})$ and $\bar{\Upsilon}^*$ are minimized for a deterministic version of SA known as quasi-stochastic approximation [22, 25], but can these ideas be extended to the stochastic setting?

- Theory in non-convergent settings, such as in non-convex optimization, will require new performance metrics.

- Although this paper focused on general single timescale SA algorithms, several algorithms of interest to the machine learning research community are built upon the framework of two time-scale SA. One example being actor-critic algorithms [20, 19]. SA in two timescales with Markovian noise has been studied before for several special cases [17, 16, 39, 11], but can we extend these results to the setting of the present paper?

References

[1] S. Allmeier and N. Gast. Computing the bias of constant-step stochastic approximation with Markovian noise. *arXiv preprint arXiv:2405.14285*, 2024.

[2] F. Bach and E. Moulines. Non-strongly-convex smooth stochastic approximation with convergence rate $o(1/n)$. In *Proc. Advances in Neural Information Processing Systems*, volume 26, pages 773–781, 2013.

[3] A. Benveniste, M. Métivier, and P. Priouret. *Adaptive algorithms and stochastic approximations*, volume 22. Springer Science & Business Media, Berlin Heidelberg, 2012.

[4] J. Bhandari, D. Russo, and R. Singal. A finite time analysis of temporal difference learning with linear function approximation. In *Conference On Learning Theory*, pages 1691–1692, 2018.

[5] V. Borkar, S. Chen, A. Devraj, I. Kontoyiannis, and S. Meyn. The ODE method for asymptotic statistics in stochastic approximation and reinforcement learning. *arXiv e-prints:2110.14427*, pages 1–50, 2021.

[6] V. S. Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Hindustan Book Agency, Delhi, India, 2nd edition, 2021.

[7] V. S. Borkar and S. P. Meyn. The ODE method for convergence of stochastic approximation and reinforcement learning. *SIAM J. Control Optim.*, 38(2):447–469, 2000.

[8] 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. *Automatica*, 146:110623, 2022.

[9] K. L. Chung. On a stochastic approximation method. *The Annals of Mathematical Statistics*, 25(3):463–483, 1954.

[10] A. M. Devraj and S. P. Meyn. Zap Q-learning. In *Proc. of the Intl. Conference on Neural Information Processing Systems*, pages 2232–2241, 2017.

[11] T. T. Doan. Finite-time analysis and restarting scheme for linear two-time-scale stochastic approximation. *SIAM Journal on Control and Optimization*, 59(4):2798–2819, 2021.

[12] A. Durmus, E. Moulines, A. Naumov, and S. Samsonov. Finite-time high-probability bounds for Polyak–Ruppert averaged iterates of linear stochastic approximation. *Mathematics of Operations Research*, 2024.

[13] P. W. Glynn and S. P. Meyn. A Liapounov bound for solutions of the Poisson equation. *Ann. Probab.*, 24(2):916–931, 1996.

[14] D. Huo, Y. Chen, and Q. Xie. Bias and extrapolation in Markovian linear stochastic approximation with constant stepsizes. *arXiv.2210.00953 (Abstract published in Proceedings of ACM SIGMETRICS , pages 81–82, 2023)*, 2022.

[15] C. Jin, P. Netrapalli, and M. Jordan. What is local optimality in nonconvex-nonconcave minimax optimization? In *International conference on machine learning*, pages 4880–4889. PMLR, 2020.

[16] M. Kaledin, E. Moulines, A. Naumov, V. Tadić, and H.-T. Wai. Finite time analysis of linear two-timescale stochastic approximation with Markovian noise. *arXiv e-prints*, page arXiv:2002.01268, Feb. 2020.

[17] P. Karmakar and S. Bhatnagar. Two time-scale stochastic approximation with controlled Markov noise and off-policy temporal-difference learning. *Math. Oper. Res.*, 43(1):130–151, 2018.

[18] D. Kingma and J. Ba. Adam: A method for stochastic optimization. *ICLR, and preprint arXiv:1412.6980*, 2015.

[19] V. Konda. *Actor-critic algorithms*. PhD thesis, Massachusetts Institute of Technology, 2002.

[20] V. R. Konda and J. N. Tsitsiklis. Actor-critic algorithms. In *Proc. Advances in Neural Information Processing Systems*, pages 1008–1014, 2000.

[21] I. Kontoyiannis and S. P. Meyn. Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electron. J. Probab.*, 10(3):61–123 (electronic), 2005.

[22] C. K. Lauand and S. Meyn. Approaching quartic convergence rates for quasi-stochastic approximation with application to gradient-free optimization. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 15743–15756. Curran Associates, Inc., 2022.

[23] C. K. Lauand and S. Meyn. Markovian foundations for quasi stochastic approximation with applications to extremum seeking control. *arXiv 2207.06371*, 2022.

[24] C. K. Lauand and S. Meyn. The curse of memory in stochastic approximation. In *Proc. IEEE Conference on Decision and Control*, pages 7803–7809, 2023.

[25] C. K. Lauand and S. Meyn. Quasi-stochastic approximation: Design principles with applications to extremum seeking control. *IEEE Control Systems Magazine*, 43(5):111–136, Oct 2023.

[26] C. J. Li, W. Mou, M. Wainwright, and M. Jordan. Root-sgd: Sharp nonasymptotics and asymptotic efficiency in a single algorithm. In *Conference on Learning Theory*, pages 909–981. PMLR, 2022.

[27] M. Metivier and P. Priouret. Theoremes de convergence presque sure pour une classe d'algorithmes stochastiques a pas decroissants. *Prob. Theory Related Fields*, 74:403–428, 1987.

[28] S. Meyn. *Control Systems and Reinforcement Learning.* Cambridge University Press, Cambridge, 2022.

[29] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability.* Cambridge University Press, Cambridge, second edition, 2009. Published in the Cambridge Mathematical Library. 1993 edition online.

[30] W. Mou, C. Junchi Li, M. J. Wainwright, P. L. Bartlett, and M. I. Jordan. On linear stochastic approximation: Fine-grained Polyak-Ruppert and non-asymptotic concentration. *Conference on Learning Theory and arXiv:2004.04719*, pages 2947–2997, 2020.

[31] B. T. Polyak. A new method of stochastic approximation type. *Avtomatika i telemekhanika (in Russian). translated in Automat. Remote Control, 51 (1991)*, pages 98–107, 1990.

[32] B. T. Polyak and A. B. Juditsky. Acceleration of stochastic approximation by averaging. *SIAM J. Control Optim.*, 30(4):838–855, 1992.

[33] D. Ruppert. Efficient estimators from a slowly convergent Robbins-Monro processes. Technical Report Tech. Rept. No. 781, Cornell University, School of Operations Research and Industrial Engineering, Ithaca, NY, 1988.

[34] R. Srikant and L. Ying. Finite-time error bounds for linear stochastic approximation and TD learning. In *Conference on Learning Theory*, pages 2803–2830, 2019.

[35] R. Sutton and A. Barto. *Reinforcement Learning: An Introduction.* MIT Press, Cambridge, MA, 2nd edition, 2018.

[36] C. Szepesvári. *Algorithms for Reinforcement Learning.* Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2010.

[37] S. Vaswani, F. Bach, and M. Schmidt. Fast and faster convergence of sgd for over-parameterized models and an accelerated perceptron. In *The 22nd international conference on artificial intelligence and statistics*, pages 1195–1204. PMLR, 2019.

[38] M. J. Wainwright. Stochastic approximation with cone-contractive operators: Sharp ℓ_∞-bounds for Q-learning. *CoRR*, abs/1905.06265, 2019.

[39] F. Zarin Faizal and V. Borkar. Functional Central Limit Theorem for Two Timescale Stochastic Approximation. *arXiv e-prints*, page arXiv:2306.05723, June 2023.

Appendix

The following contains proofs of the main results and further details on numerical experiments.

More Notation: Recall that π is the invariant measure for the Markov chain $\mathbf{\Phi}$. For a measurable function g on X we often write $\bar{g}$ or $\pi(g)$ for the mean $\int g(x)\pi(x)$. The function g may be real- or vector-valued.

For any measurable function $w\colon \mathsf{X} \to [1,\infty)$, let L_w denote the set of all measurable functions $g\colon \mathsf{X} \to \mathbb{R}$ satisfying

$$\|g\|_w := \sup_{x\in\mathsf{X}} \frac{1}{w(x)} |g(x)| < \infty$$

If $\pi(w)$ is finite then so is $\pi(|g|)$, and in this case we denote

$$\tilde{g}(x) = g(x) - \bar{g}\,, \qquad x \in \mathsf{X}\,.$$

The function $\hat{g}$ is said to be the solution to *Poisson's equation* with *forcing function g* if it satisfies

$$\mathsf{E}[\hat{g}(\Phi_{n+1}) - \hat{g}(\Phi_n) \mid \Phi_n = x] = -\tilde{g}(x)\,, \qquad x \in \mathsf{X} \tag{25}$$

For a d-dimensional vector-valued random variable X and $p \geq 1$, the L_p norm is denoted $\|X\|_p = (\mathsf{E}[\|X\|^p])^{1/p}$, and the L_p *span norm* $\|X\|_{p,s} = \min\{\|X - c\|_p : c \in \mathbb{R}^d\}$. When $p = 2$ we have

$$\|X\|_{2,s} = \sqrt{\mathrm{tr}\,(\mathsf{Cov}(X))} \tag{26a}$$

$$\|X\|_2^2 = \|X\|_{2,s}^2 + \|\mathsf{E}[X]\|^2 \tag{26b}$$

Any $d \times d$ positive definite matrix M defines a norm on $\mathbb{R}^d$ via $\|x\|_{\mathsf{M}}^2 = x^{\intercal}\mathsf{M}x$ for $x \in \mathbb{R}^d$

A Noise Decomposition for General SA

The following results may be found in [29]: Part (i) of the following follows from Theorem 16.0.1, and (ii) follows from Theorem 17.4.2. Part (iii) follows from Theorem 17.5.3 and the representation in (28) is (17.47) of Section 17.4.3.

Theorem A.1. *Suppose that $\mathbf{\Phi} := \{\Phi_n : n \geq 0\}$ is an aperiodic, ψ-irreducible Markov chain satisfying Assumption (A4). Then,*

(i) *$\mathbf{\Phi}$ is v-uniformly ergodic with $v = e^V$: there is $\varrho_v \in (0,1)$ and $b_v < \infty$ such that for any $g \in L_v$,*

$$\big|\mathsf{E}_x[\tilde{g}(\Phi_k)]\big| \leq b_v \|g\|_v v(x) \varrho_v^k$$

where the subscript x on the left hand side indicates that $\Phi_0 = x$.

(ii) *If $\|g\|_w < \infty$ with $w = v^\delta$ and $\delta \leq 1$, then there is a solution to Poisson's equation* (25)*. Moreover, $\|\hat{g}\|_w < \infty$ and the solution can be chosen so that $\pi(\hat{g}) = 0$.*

(iii) *If $g\colon \mathsf{X} \to \mathbb{R}^m$ satisfies $g_i^2 \in L_v$ for each i, then the m-dimensional stochastic process $\{Z_n^g = n^{-1/2}\sum_{k=1}^n \tilde{g}(\Phi_k) : n \geq 1\}$ converges in distribution to a Gaussian $N(0, \Sigma_{\mathsf{CLT}}^g)$ random variable, and the second moment also converges:*

$$\lim_{n\to\infty} \mathsf{E}[Z_n^g (Z_n^g)^{\intercal}] = \lim_{n\to\infty} \frac{1}{n} \mathsf{Cov}\Big(\sum_{k=1}^n g(\Phi_k)\Big) = \Sigma_{\mathsf{CLT}}^g \tag{27}$$

The $m \times m$ asymptotic covariance may be expressed in two equivalent forms:

- *the sum of the auto-covariance matrices:*

$$\Sigma_{\mathsf{CLT}}^g := \sum_{k=-\infty}^{\infty} \mathsf{E}_\pi[\tilde{g}(\Phi_0)\tilde{g}(\Phi_k)^{\intercal}]$$

in which $\{\Phi_k : -\infty < k < \infty\}$ is a stationary version of the Markov chain.

- *In terms of solutions to Poisson's equation* (25)*:*

$$\Sigma^g_{\mathsf{CLT}} = \mathsf{E}_\pi[\hat g(\Phi_0)\tilde g(\Phi_0)^\intercal] + \mathsf{E}_\pi[\tilde g(\Phi_0)\hat g(\Phi_0)^\intercal] - \mathsf{E}_\pi[\tilde g(\Phi_0)\tilde g(\Phi_0)^\intercal] \tag{28}$$

$\square$

Poisson's equation (25) is the key workhorse in establishing any of the main results in the present paper. However, we require a slight extension of the definition (25): we are interested in solutions with forcing functions of the form $h : \mathbb{R}^d \times \mathsf{X} \to \mathbb{R}^d$. Solutions for the joint process $\boldsymbol{\Psi}$ will not be required, but instead we obtain solutions for each θ. The following abuse of notation will be employed throughout the paper: $\hat h$ is said to be the solution to Poisson's equation with forcing function h if it satisfies, for each fixed $\theta \in \mathbb{R}^d$,

$$\mathsf{E}[\hat h(\theta, \Phi_{n+1}) - \hat h(\theta, \Phi_n) \mid \Phi_n = x] = -h(\theta, x) + \overline{h}(\theta)\,, \quad x \in \mathsf{X} \tag{29}$$

This notation is also extended to matrix-valued functions.

The first important example is $h = f$ that is part of the definition of the SA recursion. The following is a companion to Thm. A.1 (iii).

Proposition A.2. *Suppose that (A2) and (A4) hold. Then, $\hat f \colon \mathbb{R}^d \times \mathsf{X} \to \mathbb{R}^d$ exists solving* (29) *with forcing function f, in which $\mathsf{E}_\pi[\hat f(\theta, \Phi_n)] = 0$ for each $\theta \in \mathbb{R}^d$, and for a constant b_f and all $\theta, \theta' \in \mathbb{R}^d$ and $x \in \mathsf{X}$:*

$$\|\hat f(\theta, x)\| \le b_f\big(1 + V(x)\big)\big[1 + \|\theta\|\big] \tag{30a}$$

$$\|\hat f(\theta, x) - \hat f(\theta', x)\| \le b_f\big(1 + V(x)\big)\|\theta - \theta'\| \tag{30b}$$

Proof. Thm. A.1 (iii) cannot be applied because the bounds on f involve W rather than $w = e^{\delta V}$.

An application of Jensen's inequality implies the following bound under (DV3):

$$\mathsf{E}\big[V(\Phi_{k+1}) \mid \Phi_k = x\big] \le V(x) - W(x) + b\mathbf{1}_C(x)\,, \qquad x \in \mathsf{X}\,.$$

The conclusion then follows from [13, Th. 2.3]. $\square$

Corollary A.3. *Suppose the assumptions of Prop. A.2 along with (A5). Then, $\widehat A \colon \mathbb{R}^{d\times d} \times \mathsf{X} \to \mathbb{R}^{d\times d}$ exists solving* (29) *with forcing function A, in which $\mathsf{E}_\pi[\widehat A(\theta, \Phi_n)] = 0$ for each $\theta \in \mathbb{R}^d$, and for a constant b_A and all $\theta, \theta' \in \mathbb{R}^d$ and $x \in \mathsf{X}$:*

$$\|\widehat A(\theta, x)\|_F \le b_A\big(1 + V(x)\big)\big[1 + \|\theta\|\big] \tag{31a}$$

$$\|\widehat A(\theta, x) - \widehat A(\theta', x)\|_F \le b_A\big(1 + V(x)\big)\|\theta - \theta'\| \tag{31b}$$

$\square$

Solutions to Poisson's equation enable "whitening" of $\boldsymbol{\Delta}$ through the noise decomposition of Métivier and Priouret [27],

Lemma A.4. *Under (A2), the representation* (13) *holds with*

$$\mathcal{W}_{k+2} := \hat f(\theta_k, \Phi_{k+2}) - \mathsf{E}[\hat f(\theta_k, \Phi_{k+2}) \mid \mathcal{F}_{k+1}] \tag{32a}$$

$$\mathcal{T}_{k+1} := \hat f(\theta_k, \Phi_{k+1}) \tag{32b}$$

$$\Upsilon_{k+2} := -\frac{1}{\alpha_{k+1}}[\hat f(\theta_{k+1}, \Phi_{k+2}) - \hat f(\theta_k, \Phi_{k+2})] \tag{32c}$$

$\square$

The sequence $\{\mathcal{W}_{n+2}\}$ is a martingale difference sequence and $\{\Upsilon_{k+2}\}$ is the major source of bias for stochastic approximation with constant step-size [24]. When the noise is AD, $\hat f$ is only a function of $\boldsymbol{\Phi}$, giving $\Upsilon \equiv 0$.

B Moment Bounds

The standard ODE approximation of an SA recursion begins with the introduction of "sampling times" for the mean flow (5): $\tau_{k+1} = \tau_k + \alpha_{k+1}$, with $\tau_0 = 0$ [3, 6]. Then a fixed time horizon $T > 0$ is specified, along with the sequence $T_{n+1} = \min\{\tau_k \colon \tau_k \geq T_n + T\}$ for $n \geq 0$, initialized with $T_0 = 0$. It follows that the interval $[T_n, T_{n+1}]$ satisfies $T \leq T_{n+1} - T_n \leq T + \bar{\alpha}$ for each n, where $\bar{\alpha} := \sup_k \alpha_k$. We let m_n denote the integer satisfying $T_n = \tau_{m_n}$, and $\{\vartheta^{(n)}_t : t \geq T_n\}$ denote the solution to the mean flow with initial condition $\vartheta^{(n)}_{T_n} = \theta_{m_n}$. Convergence theory of SA proceeds by comparing $\{\theta_k : m_n \leq k \leq m_{n+1}\}$ and $\{\vartheta^{(n)}_{\tau_k} : m_n \leq k \leq m_{n+1}\}$.

We establish moment bounds of the form (10) by extending the approach of [5], based on a Lyapunov function of the form $\mathcal{V}(\theta, x) = \mathcal{V}^\circ(\theta) + \beta \mathcal{V}^\circ(\theta) v_+(x)$ in which $\beta > 0$, and

$$\mathcal{V}^\circ(\theta) = 1 + \|\theta\|^p \,, \quad v_+(x) = \mathsf{E}[\exp(V(\Phi_{m_n+1})) \mid \Phi_{m_n} = x] \,.$$

A Lyapunov bound is obtained in the following:

Proposition B.1. *Suppose the assumptions of Thm. 2.2 hold. Then for each $\rho \in (0,1)$ and $p \geq 1$ there are constants $\varrho_0 < 1$, b_ν, and $n_0 \geq 1$ such that*

$$\mathsf{E}[\mathcal{V}(\theta_{m_{n+1}}, \Phi_{m_{n+1}}) \mid \Phi_{m_n}] \leq \varrho_0 \mathcal{V}(\theta_{m_n}, \Phi_{m_n}) + b_\nu \,, \quad n \geq n_0 \tag{33}$$

□

A similar result appears in [5] with $\rho \in (1/2, 1)$ and $p = 4$. In the following we explain how these assumptions may be relaxed.

We first note that (A4) is far stronger than what is assumed in [5], in which (18) is relaxed to $\delta_L := |L|_W := \sup_x |L(x)|/W(x) < \infty$, and δ_L is assumed sufficiently small.

The assumption that δ_L is small may be gauranteed under (A4) by applying (18):

Lemma B.2. *If (A4) holds, then for any $\delta_0 > 0$ there is $r < \infty$ such that (A4) holds with (V, W_r) using $W_r(x) := \max\{r, W(x)\}$, and*

$$\delta_L(r) := \sup_x \frac{1}{W_r(x)} |L(x)| \leq \delta_0$$

Throughout the Appendix, for given $\rho \in (0,1)$ we fix $r \geq 1$, $T > 0$, and an integer p satisfying,

$$\|\vartheta_t - \theta^*\| \leq \tfrac{1}{2} \|\vartheta_0 - \theta^*\| \,, \qquad t \geq T \,,\ \vartheta_0 \in \mathbb{R}^d \tag{34a}$$

$$\max(4, 2/\rho) \leq p < p^{\max} := \frac{1}{\delta_L(r)[T+1]} \tag{34b}$$

where in (34a), $\{\vartheta_t : t \geq 0\}$ is a solution to the mean flow, and the bound holds for any initial condition ϑ_0.

Since p, r are fixed we henceforth write δ_L (suppressing dependency on r).

The following is a simple corollary to [5, Prop. 12 (ii)], which considered only $p = 4$.

Lemma B.3. *Suppose (A1)–(A4) hold. Then, the following bound holds for $m_n < k \leq m_{n+1}$:*

$$(1 + \|\theta_k\|)^p \leq 2^p \exp\Bigl(p \delta_L \sum_{j=m_n+1}^{m_{n+1}} \alpha_j W(\Phi_j)\Bigr) (1 + \|\theta_{m_n}\|)^p \tag{35}$$

□

Bounds on the conditional expectation of the right hand side of (35) are obtained by applying Lemma B.4 that follows. The proof follows from Lemma B.2, combined with an obvious extension of [5, Prop. 3].

Lemma B.4. *Under (A4) the following holds: for any n, any non-negative sequence $\{\delta_k : 1 \leq k \leq n-1\}$ satisfying $\sum \delta_k \leq 1$ there is $b_v < \infty$ and $\delta_v > 0$ such that*

$$\mathsf{E}_x \Bigl[\exp\Bigl(V(\Phi_n) + p \delta_L T \sum_{k=0}^{n-1} \delta_k W(\Phi_k) \Bigr) \Bigr] \leq b_v \exp\bigl(V(x) - \delta_v W(x) \bigr) \,, \qquad x \in \mathsf{X} \tag{36}$$

The two lemmas combined with the Markov property imply the following under our standing assumption (34b):

Lemma B.5. *Under the assumptions of Lemma B.3, the following holds for some $\delta_v > 0$, all $n \geq 0$, and $m_n < k \leq m_{n+1}$:*

$$\mathsf{E}[\exp\bigl(V(\Phi_{k+1})\bigr)(1+\|\theta_k\|)^p \mid \Phi_{m_n+1}] \leq 2^p b_v (1+\|\theta_{m_n}\|)^p \exp\bigl(V(\Phi_{m_n+1}) - \delta_v W(\Phi_{m_n})\bigr)$$

□

[5, Lemma 11] establishes the following for $p = 4$. The simple proof is easily extended to general $p \geq 4$ under (A4).

Lemma B.6. *Under (A1)–(A4) there exists constants $0 < \varrho^{\text{B.6}} < 1$, $b^{\text{B.6}} < \infty$, and a deterministic and vanishing sequence $\{\mathcal{E}_n^{\text{B.6}} : n \geq n_g\}$ such that for all $n \geq n_g$,*

$$\mathsf{E}\bigl[\bigl(\|\theta_{m_{n+1}}\| + 1\bigr)^p \mid \Phi_{m_n+1}\bigr] \leq \Bigl(\varrho^{\text{B.6}} + \mathcal{E}_n^{\text{B.6}} v(\Phi_{m_n+1})\Bigr)\bigl(1+\|\theta_{m_n}\|\bigr)^p + b^{\text{B.6}}. \tag{37}$$

Proof of Prop. B.1. Lemma B.6 bounds the first term in the Lyapunov function $\mathcal{V}(\theta,x) = \mathcal{V}^\circ(\theta) + \beta\mathcal{V}^\circ(\theta)v_+(x)$

$$\begin{aligned}\mathsf{E}[\mathcal{V}^\circ(\theta_{m_{n+1}}) \mid \Phi_{m_n}] &\leq \Bigl(\varrho^{\text{B.6}} + \mathcal{E}_n^{\text{B.6}} v_+(\Phi_{m_n})\Bigr)\bigl(1+\|\theta_{m_n}\|\bigr)^p + b^{\text{B.6}} \\ &\leq \bigl[\varrho^{\text{B.1}} + b_2^{\text{B.1}}\mathcal{E}_n^{\text{B.6}} v_+(\Phi_{m_n})\bigr]\mathcal{V}^\circ(\theta_{m_n}) + b^{\text{B.6}} + b_1^{\text{B.1}}\,,\end{aligned} \tag{38a}$$

where $\varrho^{\text{B.1}} \in (\varrho^{\text{B.6}}, 1)$ is fixed, and then $b_1^{\text{B.1}} = \sup\{\mathcal{V}^\circ(\theta) : \varrho^{\text{B.6}}(1+\|\theta\|)^p \geq \varrho^{\text{B.1}}(1+\|\theta\|^p)\}$. The second constant is the upper bound $b_2^{\text{B.1}} = \sup\{(1+x)^p(1+x^p)^{-1} : x \geq 0\}$.

To bound the second term in $\mathcal{V}$ we first apply the Markov property,

$$\mathsf{E}[\mathcal{V}^\circ(\theta_{m_{n+1}})v_+(\Phi_{m_{n+1}}) \mid \Phi_{m_n}] = \mathsf{E}[\mathcal{V}^\circ(\theta_{m_{n+1}})v(\Phi_{m_{n+1}+1}) \mid \Phi_{m_n}]$$

where $v = e^V$. Lemma B.5 then gives a bound on the conditional mean of $\mathcal{V}(\theta,x)$: for some finite constant $b_3^{\text{B.1}}$,

$$\begin{aligned}\mathsf{E}[\mathcal{V}^\circ(\theta_{m_{n+1}})v_+(\Phi_{m_{n+1}})) \mid \Phi_{m_n}] &\leq 2^p b_v (1+\|\theta_{m_n}\|)^p v_+(\Phi_{m_n}) \exp\bigl(-\delta_v W(\Phi_{m_n})\bigr) \\ &\leq 2^p b_v b_2^{\text{B.1}} \mathcal{V}^\circ(\theta_{m_n}) v_+(\Phi_{m_n}) \exp\bigl(-\delta_v W(\Phi_{m_n})\bigr) \\ &\leq \tfrac{1}{2}\mathcal{V}^\circ(\theta_{m_n})v_+(\Phi_{m_n}) + b_3^{\text{B.1}}\mathcal{V}^\circ(\theta_{m_n})\end{aligned} \tag{38b}$$

The proof of the proposition is completed on combining the bounds (38a), (38b) and choosing $\beta > 0$ so that $\varrho_1 := \beta b_3^{\text{B.1}} + \varrho^{\text{B.1}} < 1$. The constant $\varrho_0 < 1$ in (33) is $\varrho_0 = \max(1/2, \varrho_1)$. □

The following results are immediate from (33):

Corollary B.7. *Under the assumptions of Thm. 2.2,*

(i) $\sup_{k\geq 0} \mathsf{E}[\|\theta_k + 1\|^p \exp(V(\Phi_{k+1}))] < \infty$

(ii) *there is $b^{\text{B.7}}$ depending upon Ψ_0 and ρ such that*

$$\sup_{k\geq 0}\|\mathcal{W}_k\|_p \leq b^{\text{B.7}}\,, \quad \sup_{k\geq 0}\|\mathcal{T}_k\|_p \leq b^{\text{B.7}}\,, \quad \sup_{k\geq 0}\|\Upsilon_k\|_p \leq b^{\text{B.7}}$$

□

Proof of parts (i) and (ii) of Thm. 2.2. With Corollary B.7 in hands, the proof of part (i) follows from the same arguments as the ones used in the proof of [5, Lemma 16].

We now establish part (ii). In view of (10), and applying (34b), we have $\sum_{n=1}^\infty \mathsf{E}[\|\tilde{\theta}_n\|^p] < \infty$, implying $\tilde{\theta}_n \xrightarrow{\text{a.s.}} 0$, by Fubini's theorem. □

C Asymptotic Statistics for General Stochastic Approximation

We begin this section with a representation for the sample-path target bias, obtained by averaging both sides of (2) and applying the noise decomposition in Lemma A.4:

Lemma C.1. *The following holds for the recursion* (2)*,*

$$\frac{1}{N}\sum_{k=1}^{N}\bar{f}(\theta_k) = \frac{1}{N}(S_N^{\tau} - S_{N+1}^{\Delta}) \tag{39a}$$

where $S_N^{\tau} = \sum_{k=1}^{N}\frac{1}{\alpha_{k+1}}(\theta_{k+1}-\theta_k)$ *and*

$$S_{N+1}^{\Delta} = \sum_{k=1}^{N}\Delta_{k+1} = M_{N+2} - \mathcal{T}_{N+2} + \mathcal{T}_2 - \sum_{k=1}^{N}\alpha_{k+1}\Upsilon_{k+2} \tag{39b}$$

in which $\{M_{N+2} := \sum_{k=1}^{N}\mathcal{W}_{k+2} : N \ge 1\}$ *is a martingale.* □

The next two lemmas will be used repeatedly to establish several of the remaining results of the paper.

Lemma C.2 (Summation by Parts)**.** *For any two real-valued sequences* $\{x_n, y_n : n \ge 0\}$ *and integers* $0 \le N_0 < N$*,*

$$\sum_{k=N_0+1}^{N} x_k(y_k - y_{k-1}) = x_{N+1}y_N - x_{N_0+1}y_{N_0} - \sum_{k=N_0+1}^{N}(x_{k+1}-x_k)y_k$$

□

Lemma C.3. *Under (A1) the following bounds hold: for a constant* $b^{\mathrm{C.3}} < \infty$

(i) $\displaystyle\sum_{k=1}^{N}\Big|\frac{1}{\alpha_{k+1}} - \frac{1}{\alpha_k}\Big|\sqrt{\alpha_k} \le \frac{2}{\sqrt{\alpha_0}}N^{\rho/2} + b^{\mathrm{C.3}}$

(ii) $\displaystyle\sum_{k=1}^{N}\alpha_k \le \frac{\alpha_0}{1-\rho}N^{1-\rho} + b^{\mathrm{C.3}}$

(iii) $\displaystyle\sum_{k=1}^{N}\frac{1}{k+1} \le \log(N+1) + b^{\mathrm{C.3}}$

□

In order to obtain moment bounds for PR averaging, Lemma C.1 tells us we need to first analyze the right hand side of (39a). Lemmas C.4 and C.5 establish moment bounds for S^{τ} as well as for each term in (39b).

Lemma C.4. *Under the assumptions of Thm. 2.2, there is* $b^{\mathrm{C.4}}$ *depending upon* Ψ_0 *and* ρ *such that the following bounds hold*

(i) $\|S_N^{\tau}\|_2 \le b^{\mathrm{C.4}}N^{\rho/2}$

(ii) $\|\mathcal{T}_{N+2} - \mathcal{T}_2\|_2 \le b^{\mathrm{C.4}}$

(iii) $\|\sum_{k=1}^{N}\alpha_{k+1}\Upsilon_{k+2}\|_p \le b^{\mathrm{C.4}}N^{1-\rho}$

Proof. To prove part (i), we apply Lemma C.2, followed by the triangle inequality to obtain

$$\begin{aligned}\|S_N^{\tau}\|_2 &= \Big\|\frac{1}{\alpha_{N+1}}\tilde{\theta}_{N+1} - \frac{1}{\alpha_2}\tilde{\theta}_1 - \sum_{k=1}^{N}\Big(\frac{1}{\alpha_{k+1}} - \frac{1}{\alpha_k}\Big)\tilde{\theta}_k\Big\|_2 \\ &\le \frac{1}{\alpha_{N+1}}\|\tilde{\theta}_{N+1}\|_2 + \frac{1}{\alpha_2}\|\tilde{\theta}_1\|_2 + \sum_{k=1}^{N}\Big|\frac{1}{\alpha_{k+1}} - \frac{1}{\alpha_k}\Big|\|\tilde{\theta}_k\|_2\end{aligned}$$

Jensen's inequality along with the bounds in Thm. 2.2 gives $\|\tilde{\theta}_n\|_2 \le \|\tilde{\theta}_n\|_p \le b^{2.2}\sqrt{\alpha_n}$. An application of part (i) of Lemma C.3 yields the final bound.

The result in part (ii) is implied directly from Corollary B.7 (ii) and the triangle inequality.

Part (iii) is obtained as follows:

$$\Big\| \sum_{k=1}^{N} \alpha_{k+1} \Upsilon_{k+2} \Big\|_p \le \Big(\sup_{k\ge 0} \|\Upsilon_{k+2}\|_p \Big) \sum_{k=1}^{N} |\alpha_{k+1}|$$

and Corollary B.7 (ii) gives $\sup_{k\ge 0} \|\Upsilon_{k+2}\|_p \le b^{\mathrm{B.7}}$. An application of part (ii) of Lemma C.3 finishes the proof. □

The following notation is adopted to define parameter independent disturbance processes: let $M^*_{N+2} = \sum_{k=1}^{N} \mathcal{W}^*_{n+2}$ with

$$\mathcal{W}^*_{n+2} := \hat{f}(\theta^*, \Phi_{k+2}) - \mathsf{E}[\hat{f}(\theta^*, \Phi_{k+2}) \mid \mathcal{F}_{k+1}] \tag{40}$$

The above martingale difference sequence is the "noise" sequence defining the optimal asymptotic covariance of SA (19).

Lemma C.5. *Under the assumptions of Thm. 2.2, there is $b^{\mathrm{C.5}}$ depending upon Ψ_0 and ρ such that the following bounds hold*

(i) $\|M_{N+2} - M^*_{N+2}\|_2 \le b^{\mathrm{C.5}} N^{(1-\rho)/2}$

(ii) $\|M^*_{N+2}\|_2 \le b^{\mathrm{C.5}} N^{1/2}$

Proof. For part (i) we use the martingale difference property to obtain

$$\|M_{N+2} - M^*_{N+2}\|_2^2 = \Big\| \sum_{k=1}^{N} \mathcal{W}_{k+2} - \mathcal{W}^*_{k+2} \Big\|_2^2 = \sum_{k=1}^{N} \|\mathcal{W}_{k+2} - \mathcal{W}^*_{k+2}\|_2^2$$

From Prop. A.2 and Thm. 2.2, we have that $\|\mathcal{W}_{k+2} - \mathcal{W}^*_{k+2}\|^2 \le b^{\mathrm{A.2}} \|\tilde{\theta}_n\|_2^2 \le b^{\mathrm{A.2}} b^{2.2} \alpha_n$ for some constant $b^{\mathrm{C.5}}$ depending upon Ψ_0. An application of part (ii) of Lemma C.3 yields (i).

The proof of part (ii) follows similarly to (i): since $\{\mathcal{W}^*_{k+2}\}$ is a martingale difference sequence,

$$\|M^*_{N+2}\|_2^2 = \sum_{k=1}^{N} \|\mathcal{W}^*_{k+2}\|_2^2 \le N \Big(\sup_{k\ge 0} \|\mathcal{W}^*_{k+2}\|_2^2 \Big) \le b^{\mathrm{B.7}} N$$

where the last bound is obtained from an application of Corollary B.7 (ii). □

We are now equipped to prove part (iii) of Thm. 2.2. Only the proof for the multiplicative (MU) noise case is given. The result for when the noise is AD follows almost identically to the MU setting with the only difference being that $\Upsilon \equiv 0$, as explained after Lemma A.4.

Proof of part (iii) of Thm. 2.2. A Taylor series expansion of $\bar{f}(\theta_k)$ around θ^* gives

$$\bar{f}(\theta_k) = \bar{f}(\theta^*) + A^*(\theta_k - \theta^*) + \mathcal{E}^T_k \tag{41a}$$

where under the assumptions of the theorem, the second order Taylor series expansion error $\mathcal{E}^T$ admits the upper bound:

$$\|\mathcal{E}^T_k\|_p \le L_A \|\tilde{\theta}_k\|_p^2 \le L_A b^{2.2} \alpha_k \tag{41b}$$

in which L_A is the Lipschitz constant associated with $\bar{A}$.

In view of (41), (39a) is equivalently expressed as

$$\theta^{\mathsf{PR}}_N - \theta^* = [A^*]^{-1} \frac{1}{N} \Big[S^{\tau}_N - S^{\Delta}_{N+1} + \sum_{k=1}^{N} \mathcal{E}^T_k + M^*_{N+2} - M^*_{N+2} \Big]$$

Taking L_2 norms on both sides and applying the triangle inequality we obtain,

$$
\begin{aligned}
\|\tilde{\theta}^{\text{PR}}_N\|_2 &\leq \|[A^*]^{-1}\|_F \frac{1}{N}\Big(\|S^{\tau}_N\|_2 + \|S^{\Delta}_{N+1} - M^*_{N+2}\|_2 + \sum_{k=1}^{N}\|\mathcal{E}^T_k\|_2 + \|M^*_{N+2}\|_2\Big) \\
&\leq \|[A^*]^{-1}\|_F \frac{1}{N}(\varepsilon^a_N + \varepsilon^b_N + \varepsilon^c_N)
\end{aligned}
$$

in which $\|\cdot\|_F$ denotes the Frobenius norm and the identity (39b) gives

$$
\begin{aligned}
\varepsilon^a_N &= \|S^{\tau}_N\|_2 + \|M_{N+2} - M^*_{N+2}\|_2 + \|\mathcal{T}_{N+2} + \mathcal{T}_2\|_2 \\
\varepsilon^b_N &= \Big\|\sum_{k=1}^{N}\alpha_{k+1}\Upsilon_{k+2}\Big\|_2 + \sum_{k=1}^{N}\|\mathcal{E}^T_k\|_2 \\
\varepsilon^c_N &= \|M^*_{N+2}\|_2
\end{aligned}
$$

Lemmas C.4 and C.5 and (41b) yield the following bounds: for a constant $b^{2.2}$ depending upon Ψ_0 and ρ,

$$
\varepsilon^a_N \leq b^{2.2}\max\{N^{\rho/2}, N^{(1-\rho)/2}\}, \quad \varepsilon^b_N \leq b^{2.2}N^{1-\rho}, \quad \varepsilon^c_N \leq b^{2.2}N^{1/2}
$$

The error is dominated by ε^b_N when $\rho < 1/2$ and by ε^c_N when $\rho > 1/2$, completing the proof. □

We conclude this section with a proof of the identity in (11) for the average target bias: let $\widehat{A}(\theta, x) = \partial_\theta \hat{f}(\theta, x)$ and

$$
\Upsilon^*_{k+2} := -\widehat{A}(\theta^*, \Phi_{k+2})f(\theta^*, \Phi_{k+1}), \quad \bar{\Upsilon}^* = \mathsf{E}[\Upsilon^*_{k+2}] \tag{42}
$$

Lemma C.6. *Suppose (A1)–(A5) hold. Then, for each $\rho \in (0,1)$ there is $b^{\text{C.6}}$ depending upon Ψ_0 and ρ,*

(i) $\Upsilon_{k+2} = -\widehat{A}(\theta_k, \Phi_{k+2})f(\theta_k, \Phi_{k+1}) + b^{\text{C.6}}\alpha_{k+1}$

(ii) $\|\Upsilon_{k+2} - \Upsilon^*_{k+2}\|_2 \leq b^{\text{C.6}}\alpha^{1/2}_{k+1}$

Proof. A Taylor series expansion of $\hat{f}$ around θ_k gives for each k

$$
\hat{f}(\theta_{k+1}, \Phi_{k+2}) - \hat{f}(\theta_k, \Phi_{k+2}) = \widehat{A}(\theta_k, \Phi_{k+2})(\theta_{k+1} - \theta_k) + \mathcal{E}^T_k \tag{43}
$$

where $\mathcal{E}^T$ denotes the approximation error. By the mean value theorem, there is $\theta^\circ \in (\theta_k, \theta_{k+1})$ such that $\mathcal{E}^T_k = [\widehat{A}(\theta^\circ, \Phi_{k+2}) - \widehat{A}(\theta_k, \Phi_{k+2})](\theta_{k+1} - \theta_k)$. Corollary A.3 implies the upper bound: $\|\mathcal{E}^{\widehat{A}}_k\|_2 \leq b^{\text{C.6}}\|\theta_{k+1} - \theta_k\|^2_2$ with $b^{\text{C.6}}$ depending upon Ψ_0 and ρ.

Using the identity (6) and Lipschitz continuity of $\bar{f}$, we obtain, through the triangle inequality,

$$
\|\theta_{k+1} - \theta_k\|_2 = \alpha_{k+1}\|\bar{f}(\theta_k) + \Delta_{k+1}\|_2 \leq \alpha_{k+1}(L_{\bar{f}}\|\tilde{\theta}_k\|_2 + \|\Delta_{k+1}\|_2) \leq \alpha_{k+1}b^{\text{C.6}}
$$

in which $L_{\bar{f}}$ is the Lipschitz constant associated with $\bar{f}$ and the last bound follows from (10) and Corollary B.7 (ii), for a potentially larger $b^{\text{C.6}}$.

Applying (2) to the right side of (43) completes the proof of part (i), in view of the definition of Υ in (32c).

Part (ii) follows from Lipschitz continuity of $\widehat{A}f$, (10) and part (i), via the triangle inequality. □

Proof of part (iv) of Thm. 2.2. The first step is to take expectations of both sides of (39a). From (13), we obtain $\mathsf{E}[\Delta_{k+1}] = \mathsf{E}[-\mathcal{T}_{k+2} + \mathcal{T}_{k+1} - \alpha_{k+1}\Upsilon_{k+2}]$, which in turn gives,

$$
\frac{1}{N}\sum_{k=1}^{N}\mathsf{E}[\bar{f}(\theta_k)] = \frac{1}{N}\Big(\mathsf{E}[S^{\tau}_N] + \mathsf{E}[\mathcal{T}_{N+2} - \mathcal{T}_2] + \sum_{k=1}^{N}\alpha_{k+1}\mathsf{E}[\Upsilon_{k+2}]\Big)
$$

Adding and subtracting $\frac{1}{N}\sum_{k=1}^{N}\alpha_{k+1}\bar{\Upsilon}^*$ to the right side of the above equation with $\bar{\Upsilon}^*$ defined in (42) yields

$$
\frac{1}{N}\sum_{k=1}^{N}\mathsf{E}[\bar{f}(\theta_k)] = \frac{1}{N}\Big(\bar{\Upsilon}^*\sum_{k=1}^{N}[\alpha_{k+1} + \varepsilon^a_k] + \varepsilon^b_N\Big) \tag{44}
$$

in which

$$\varepsilon_k^a = \alpha_{k+1}\mathsf{E}[\Upsilon_{k+2} - \Upsilon_{k+2}^*]$$
$$\varepsilon_N^b = \mathsf{E}[S_N^\tau] + \mathsf{E}[\mathcal{T}_{N+2} - \mathcal{T}_2]$$

Bounds on the terms constituting ε^b are given by Lemma C.4 and are as follows:

$$\begin{aligned}\|\mathsf{E}[S_N^\tau]\| &\le \|S_N^\tau\|_2 \le b^{\mathrm{C}.4}N^{\rho/2}\\ \|\mathsf{E}[\mathcal{T}_{N+2} - \mathcal{T}_2]\| &\le \|\mathcal{T}_{N+2} - \mathcal{T}_2\|_2 \le b^{\mathrm{C}.4}\end{aligned} \tag{45}$$

Consequently, $\|\varepsilon_N^b\| \le b^{\mathrm{C}.4}N^{\rho/2}$. A bound on ε^a is achieved from Lemma C.6:

$$\|\varepsilon_k^a\| \le |\alpha_{k+1}|\|\mathsf{E}[\Upsilon_{k+2} - \Upsilon_{k+2}^*]\| \le |\alpha_{k+1}|\|\Upsilon_{k+2} - \Upsilon_{k+2}^*\|_2 \le b^{\mathrm{C}.6}\alpha_{k+1}^{3/2}$$

where the constants $b^{\mathrm{C}.4}$ and $b^{\mathrm{C}.6}$ depend upon Ψ_0 and ρ.

Using the upper bounds for ε^a and ε^b, we apply Lemma C.3 (ii) to the right side of (44) and divide both sides by α_{N+1} to obtain

$$\frac{1}{\alpha_{N+1}}\sum_{k=1}^{N}\frac{1}{N}\mathsf{E}[\bar{f}(\theta_k)] = \frac{1}{1-\rho}\bar{\Upsilon}^* + \varepsilon_N^c \tag{46}$$

in which

$$\|\varepsilon_N^c\| \le b^{2.4}\max\{N^{-\rho/2}, N^{3\rho/2-1}\}$$

where $b^{2.4}$ is a constant depending upon Ψ_0 and ρ. The norm of ε_N^c is dominated by ε_N^a for $\rho < 1/2$ and by ε_N^b for $\rho > 1/2$. Moreover, ε_N^b also dominates the target bias for $\rho > 2/3$.

Taking limits of both sides of the above equation yields (11) for $\rho \in (0, 2/3)$.

□

Proof of Corollary 2.3. Results for the MD case are immediate from analysis in the AD case since the noise decomposition in Lemma A.4 is not needed: when the noise is MD, $\{\frac{1}{N}S_N^\Delta\}$ is analyzed using the same arguments as the ones used for the stochastic process $\{\frac{1}{N}M_{N+2}\}$ in the AD case. □

D Noise Decomposition for Linear SA

In general, the representation for $\boldsymbol{\Delta}$ in Lemma A.4 does not appear to be enough to obtain the asymptotic covariance for recursions with multiplicative noise, and $\rho < 1$. Restricting to linear recursions allows us to obtain the finer representation for $\boldsymbol{\Delta}$ in (14). This expression is obtained based on recursive decompositions of stochastic processes that are affine in the parameter. One example is $\Upsilon_{k+2} = -\widehat{A}_{k+2}A_{k+1}\theta_k + \widehat{A}_{k+2}b_{k+1}$, $k \ge 0$ (recall (55a)).

The next two lemmas define the main steps in each recursion. The first considers a generalization of this example, a stochastic process of the form $\mathcal{G}_{k+\ell}^+ = \mathcal{M}_{k+\ell}\theta_k + \gamma_{k+\ell}$ with matrix-valued stochastic process of the form $\mathcal{M}_{k+\ell} = \mathcal{M}(\Phi_{k+1}^{k+\ell})$ and vector-valued stochastic process $\gamma_{k+\ell} = \gamma(\Phi_{k+1}^{k+\ell})$. In the special case of Υ_{k+2} we have $\ell = 2$.

Lemma D.1. *Consider the stochastic process $\mathcal{G}_{k+\ell}^+ = \mathcal{M}_{k+\ell}\theta_k + \gamma_{k+\ell}$, with $\mathcal{M}_{k+\ell} = \mathcal{M}(\Phi_{k+1}^{k+\ell})$ and $\gamma_{k+\ell} = \gamma(\Phi_{k+1}^{k+\ell})$. Suppose that $\mathcal{M}\colon \mathsf{X}^\ell \to \mathbb{R}^{d\times d}$ and $\gamma : \mathsf{X}^\ell \to \mathbb{R}^d$ satisfy for some $b_{\mathcal{G}^+} < \infty$ and all $x_1^\ell \in \mathsf{X}^\ell$,*

$$\mathsf{E}[\|\mathcal{M}(\Phi_{k+1}^{k+\ell})\|_F^p + \|\gamma(\Phi_{k+1}^{k+\ell})\|^p \mid \Phi_k = x] \le b_{\mathcal{G}^+}v_+(x)$$

Then,

$$\mathcal{G}_{k+\ell}^+ = \mathcal{W}_{k+\ell}^{\mathcal{G}^+} + M^+(\Phi_{k+1})\theta_k + u^+(\Phi_{k+1})\,,$$

in which $\|M^+(x)\|_F + \|u^+(x)\| \le b^{\mathrm{D}.1}v_+^{1/p}(x)$ for any x, and $\{\mathcal{W}_{k+\ell}^{\mathcal{G}^+}\}$ is a martingale difference sequence satisfying

$$\mathsf{E}[\|\mathcal{W}_{k+\ell}^{\mathcal{G}^+}\|_F^p \mid \Psi_k = z] \le b^{\mathrm{D}.1}\mathcal{V}(z)$$

Proof. Let $M^+(x)\theta + u^+(x) = \mathsf{E}[\mathcal{G}^+(\theta_k, \Phi_{k+1}^{k+\ell}) \mid \Psi_k = z]$. Then,

$$\mathcal{G}_{k+\ell}^+ = \mathcal{W}_{k+\ell}^{\mathcal{G}^+} + M^+(\Phi_{k+1})\theta_k + u^+(\Phi_{k+1})$$

with $\mathcal{W}_{k+\ell}^{\mathcal{G}^+} = \mathcal{G}_{k+\ell}^+ - M^+(\Phi_{k+1})\theta_k - u^+(\Phi_{k+1})$. □

The next result is more subtle as it relies on solutions to Poisson's equation as in Lemma A.4. Its proof is identical to the proof of (13).

Lemma D.2. *Suppose the matrix valued function* $M\colon \mathsf{X} \to \mathbb{R}^{d\times d}$ *and vector valued function* $u\colon \mathsf{X} \to \mathbb{R}^d$ *defining* $G(\Psi_k) = M(\Phi_{k+1})\theta_k + u(\Phi_{k+1})$ *satisfy* $\|M(x)\|_F + \|u(x)\| \le b^{\text{D.2}} v_+^{1/p}(x)$*, with respective means* $\overline{M}$*,* $\bar{u}$*, and let* $\widehat{M}\colon \mathsf{X} \to \mathbb{R}^{d\times d}$ *and* $\hat{u}\colon \mathsf{X} \to \mathbb{R}^d$ *denote zero-mean solutions to Poisson's equation* (25)*:*

$$\mathsf{E}[\widehat{M}(\Phi_{n+1}) - \widehat{M}(\Phi_n) \mid \Phi_n = x] = M(x) - \overline{M} \qquad \mathsf{E}[\hat{u}(\Phi_{n+1}) - \hat{u}(\Phi_n) \mid \Phi_n = x] = u(x) - \bar{u}$$

Then, denoting $\widehat{G}(\theta, x) = \widehat{M}(x)\theta + \hat{u}(x)$*, we have for* $k \ge 0$*,*

$$\begin{aligned} G(\Psi_k) &= M(\Phi_{k+1})\theta_k + u(\Phi_{k+1}) \\ &= \mathcal{W}^G_{k+2} - \mathcal{T}^G_{k+2} + \mathcal{T}^G_{k+1} - \alpha_{k+1}\Upsilon^G_{k+2} + \bar{G}(\theta_k) \end{aligned} \tag{47}$$

where $\bar{G}(\theta_k) = \overline{M}\theta_k + \bar{u}$*, and for a constant* $b^{\text{D.2}}$ *independent of* Ψ_0*,*

- $\{\mathcal{W}^G_{k+2} := \widehat{G}(\theta_k, \Phi_{k+2}) - \mathsf{E}[\widehat{G}(\theta_k, \Phi_{k+2}) \mid \mathcal{F}_{k+1}]\}$ *is a martingale difference sequence satisfying* $\mathsf{E}[\|\mathcal{W}^G_{k+2}\|^p \mid \Psi_k = z] \le b^{\text{D.2}}\mathcal{V}(z)$*.*
- $\mathcal{T}^G_{k+1} := \widehat{G}(\theta_k, \Phi_{k+1})$ *with* $\mathsf{E}[\|\mathcal{T}^G_{k+1}\|^p \mid \Psi_k = z] \le b^{\text{D.2}}\mathcal{V}(z)$*.*
- $\Upsilon^G_{k+2} := -\dfrac{1}{\alpha_{n+1}}[\widehat{G}(\theta_{k+1}, \Phi_{k+2}) - \widehat{G}(\theta_k, \Phi_{k+2})]$ *with* $\mathsf{E}[\|\Upsilon^G_{k+2}\|^p \mid \Psi_k = z] \le b^{\text{D.2}}\mathcal{V}(z)$*.* □

Suppose we are in the setting of Lemma D.2 with $G(\Psi_k) = M(\Phi_{k+1})\theta_k + u(\Phi_{k+1})$ represented via (47), and denote $\mathcal{G}^+_{k+2} = \Upsilon^G_{k+2}$. We then apply Lemma D.1 to this stochastic process to obtain

$$\mathcal{G}^+_{k+2} = \mathcal{W}^{\mathcal{G}^+}_{k+2} + G^+(\Phi_{k+1})\,, \qquad G^+(\Phi_{k+1}) := M^+(\Phi_{k+1})\theta_k + u^+(\Phi_{k+1})\,.$$

Consequently, the previous two lemmas imply the following:

Proposition D.3. *Suppose the matrix valued function* $M\colon \mathsf{X} \to \mathbb{R}^{d\times d}$ *and vector valued function* $u\colon \mathsf{X} \to \mathbb{R}^d$ *satisfy* $\|M(x)\|_F + \|u(x)\| \le b^{\text{D.2}} v_+^{1/p}(x)$*. Then,*

$$\begin{aligned} G(\Psi_k) &= M(\Phi_{k+1})\theta_k + u(\Phi_{k+1}) \\ &= \mathcal{D}^G_{k+2} - \mathcal{T}^G_{k+2} + \mathcal{T}^G_{k+1} - \alpha_{k+1}G^+(\Psi_k) + \bar{G}(\theta_k) \end{aligned} \tag{48}$$

where for a constant $b^{\text{D.3}}$ *independent of* Ψ_0*,*

- $\{\mathcal{D}^G_{k+2} := \mathcal{W}^G_{k+2} - \alpha_{k+1}\mathcal{W}^{\mathcal{G}^+}_{k+2}\}$ *is a martingale difference sequence satisfying* $\mathsf{E}[\|\mathcal{D}^G_{k+2}\|^p \mid \Psi_k = z] \le b^{\text{D.3}}\mathcal{V}(z)$*.*
- *There is a matrix valued function* $M^+\colon \mathsf{X} \to \mathbb{R}^{d\times d}$ *and vector valued function* $u^+\colon \mathsf{X} \to \mathbb{R}^d$*, with respective steady-state means* $\overline{M}^+$*,* $\bar{u}^+$ *such that*

$$\begin{aligned} &G^+(\Psi_k) = M^+(\Phi_{k+1})\theta_k + u^+(\Phi_{k+1}) \\ &\bar{G}^+(\theta_k) = \overline{M}^+\theta_k + \bar{u}^+ \\ &\textit{with} \quad \|M^+(x)\|_F + \|\overline{M}^+\|_F + \|u^+(x)\| + \|\bar{u}^+\| \le b^{\text{D.3}} v_+^{1/p}(x) \end{aligned} \tag{49}$$

Construction of the disturbance decomposition The representation in (14) is obtained by applying Prop. D.3 recursively, to construct a sequence of stochastic processes of the form,

$$G^{(i)}(\Psi_k) = M^{(i)}(\Phi_{k+1})\theta_k + u^{(i)}(\Phi_{k+1}) \tag{50a}$$

$$\qquad = \mathcal{D}^{(i)}_{k+2} - \mathcal{T}^{(i)}_{k+2} + \mathcal{T}^{(i)}_{k+1} - \alpha_{k+1}G^{(i+1)}(\Psi_k) + \bar{G}^{(i)}(\theta_k) \tag{50b}$$

where $\mathcal{D}^{(i)}_{k+2} := \mathcal{W}^{(i)}_{k+2} - \alpha_{k+1}\mathcal{W}^{\Upsilon^{(i)}}_{k+2}$. For each i, we denote

$$\mathcal{W}^{(i)}_{k+2} = \mathcal{W}^G_{k+2}\,, \quad \mathcal{T}^{(i)}_{k+1} = \mathcal{T}^G_{k+1}\,, \quad \Upsilon^{(i)}_{k+2} := \Upsilon^G_{k+2}\,, \quad \mathcal{W}^{\Upsilon^{(i)}}_{k+2} := \mathcal{W}^{\Upsilon^G}_{k+2} \tag{50c}$$

in which these terms are defined as in Lemmas D.1 and D.2 with $G(\Psi_k) = G^{(i)}(\Psi_k)$ in application to the first lemma and $\mathcal{G}_{k+2} = \Upsilon^G_{k+2}$ in application to the second.

The recursion is initialized using

$$
\begin{aligned}
G^{(0)}(\Psi_k) &= f(\Psi_k) = M^{(0)}(\Phi_{k+1})\theta_k + u^{(0)}(\Phi_{k+1}) \\
\bar{G}^{(0)}(\theta_k) &= \bar{f}(\theta_k) = \overline{M}^{(0)}\theta_k + \bar{u}^{(0)} \\
\text{where } M^{(0)}(\Phi_{k+1}) &= A_{k+1}\,, \quad u^{(0)}(\Phi_{k+1}) = -b_{k+1} \\
\overline{M}^{(0)} &= A^*\,, \qquad\qquad \bar{u}^{(0)} = -\bar{b}
\end{aligned}
$$

The subsequent functions composing the sequence of affine functions $\{G^{(i)}\}$ in (50) are defined through the following steps:

- For $1 \le i \le m-1$: Given $M^{(i)}(\Phi_{k+1})$ and $u^{(i)}(\Phi_{k+1})$ that define $G^{(i)}(\Psi_k)$ via (50a), apply Prop. D.3 with $G(\Psi_k) = G^{(i)}(\Psi_k)$ to obtain the terms $\{\mathcal{W}_{k+2}^{(i)}, \mathcal{T}_{k+2}^{(i)}, \mathcal{W}_{k+2}^{\Upsilon^{(i)}}, G^{(i+1)}(\Psi_k)\}$ in (50b). In particular, the pair $\{M^+(\Phi_{k+1}), u^+(\Phi_{k+1})\}$ resulting from Prop. D.3 defines $\{M^{(i+1)}(\Phi_{k+1}), u^{(i+1)}(\Phi_{k+1})\}$.
- For $i = m$, Lemma D.2 is applied to (50a) one final time to generate $\{\mathcal{W}_{k+2}^{(m)}, \mathcal{T}_{k+2}^{(m)}, \Upsilon_{k+2}^{(m)}, \mathcal{W}_{k+2}^{\Upsilon^{(m)}}\}$ in (50c).

Proposition D.4. *Subject to the assumptions of Thm. 2.4, and for any fixed $m \ge 1$, the representation* (14) *holds for each $k \ge 0$. Moreover, the terms in* (14) *satisfy the following for a constant $b^{\mathrm{D.4}}$ depending upon Ψ_0 and ρ:*

(i) $\sup_k \|\Upsilon_{k+2}^\bullet\|_2 \le b^{\mathrm{D.4}}$

(ii) $\{\mathcal{W}_{k+2}^\bullet : k \ge 0\}$ *is a martingale difference sequence satisfying* $\sup_k \|\mathcal{W}_{k+2}^\bullet\|_2 \le b^{\mathrm{D.4}}$

(iii) $\sup_k \|\mathcal{T}_{k+1}^\bullet\|_2 \le b^{\mathrm{D.4}}$*, and* $\|\mathcal{T}_{k+1}^\bullet - \mathcal{R}_{k+1}^\bullet\|_2 \le b^{\mathrm{D.4}}\alpha_{k+1}/k$ *for* $k \ge 0$

(iv) *The deterministic sequences of matrices $\{\bar{G}_{k+1}\}$ and vectors $\{\bar{u}_{k+1}\}$ are convergent, and hence uniformly bounded in k. Moreover,*

$$
\|\alpha_{k+1}\bar{G}_{k+1} - \alpha_k\bar{G}_k\|_F \le b^{\mathrm{D.4}}\alpha_k/k\,, \qquad k \ge 1
$$

The special case $m = 0$ yields (13) *and $\mathcal{W}_{k+2}^{(0)} = \mathcal{W}_{k+2}$, $\mathcal{T}_{k+2}^{(0)} = \mathcal{T}_{k+2}$ in which $\{\mathcal{W}_k\}$ and $\{\mathcal{T}_k\}$ are defined in* (32).

Proof. Following the steps outlined after (50b), we obtain (14) in which,

$$
\begin{aligned}
\mathcal{W}_{k+2}^\bullet &= -\beta_{k+1}^{m+1}\mathcal{W}_{k+2}^{\Upsilon^{(m)}} + \sum_{i=0}^m \beta_{k+1}^i \mathcal{D}_{k+2}^{(i)} & \Upsilon_{k+2}^\bullet &= \Upsilon_{k+2}^{(m)} \\
\alpha_{k+1}\bar{G}_{k+1}\tilde{\theta}_k &= \sum_{i=1}^m \beta_{k+1}^i \overline{M}^{(i)}\tilde{\theta}_k & \mathcal{T}_{k+2}^\bullet &= \sum_{i=0}^m \beta_{k+1}^i \mathcal{T}_{k+2}^{(i)} \\
\alpha_{k+1}\bar{u}_{k+1} &= -\sum_{i=1}^m \beta_{k+1}^i [\bar{u}^i + \overline{M}^{(i)}\theta^*] & \mathcal{R}_{k+1}^\bullet &= \mathcal{T}_{k+1}^\bullet + \sum_{i=0}^m (\beta_{k+1}^i - \beta_k^i)\mathcal{T}_{k+1}^{(i)}
\end{aligned}
\tag{51}
$$

where $\beta_k^i = (-\alpha_k)^i$, $\{\mathcal{D}_{k+2}^{(i)}\}$ is a martingale difference sequence as defined after (50b) and $\overline{M}^{(i)}, \bar{u}^{(i)}$ define $\bar{G}^{(i)}(\theta) = \overline{M}^{(i)}\theta + \bar{u}^{(i)}$ for each $\theta \in \mathbb{R}^d$ and i.

The moment bounds in parts (i)–(iii) are immediate from the bounds given as results of Lemma D.2 and Prop. D.3. The coupling between $\mathcal{T}_k^\bullet$ and $\mathcal{R}_k^\bullet$ in part (iii) is a consequence of the definitions in (51) and the triangle inequality: for a constant $b^{\mathrm{D.4}}$ depending upon Ψ_0 and ρ,

$$
\|\mathcal{R}_{k+1}^\bullet - \mathcal{T}_{k+1}^\bullet\|_2 \le \sum_{i=1}^m |\beta_{k+1}^i - \beta_k^i| \|\mathcal{T}_{k+1}^{(i)}\|_2 \le b^{\mathrm{D.4}} N^{-1-\rho}
$$

The last bound holds since the term associated with $i = 1$ dominates.

It remains to prove (iv). For each θ, the following is immediate from the definitions in (51): $\lim_{k\to\infty} \bar{G}_{k+1}[\theta - \theta^*] - \bar{u}_{k+1} = -\bar{G}^{(1)}(\theta)$. Moreover, by the triangle inequality,

$$
\|\alpha_{k+1}\bar{G}_{k+1} - \alpha_k\bar{G}_k\|_F \le \sum_{i=1}^m |\beta_{k+1}^i - \beta_k^i| \|\overline{M}^{(i)}\|_F \le b^{\mathrm{D.4}} N^{-1-\rho}
$$

□

Analogously to the definition of $\{\mathcal{W}_k^*\}$ in (40), we adopt the notation $\{\mathcal{W}_k^{\bullet *}, \mathcal{D}_{k+2}^{(i)*}, \mathcal{W}_{k+2}^{\Upsilon^{(m)*}}\}$ to define parameter independent disturbance processes. Moreover, we denote

$$H_{N+2} = \sum_{k=1}^{N} \mathcal{W}_{k+2}^{\bullet}\,, \quad H_{N+2}^{*} = \sum_{k=1}^{N} \mathcal{W}_{k+2}^{\bullet *} \tag{52}$$

We conclude this subsection by obtaining moment bounds on the terms defining $\{\mathcal{W}_{k+2}^{\bullet}\}$ in (51).

Lemma D.5. *Suppose that the assumptions of Prop. D.4 hold. Then, if for fixed $\rho \in (0,1)$, if m in (14) is chosen such that $m\rho > 1$, then there exists $b^{\text{D.5}}$ depending upon Ψ_0 and ρ such that the following bounds hold: for $0 \le i \le m$:*

(i) $\| \sum_{k=1}^N \beta_{k+1}^i (\mathcal{D}_{k+2}^{(i)} - \mathcal{D}_{k+2}^{(i)*})\|_2 \le b^{\text{D.5}} \max\{N^{(1-[2i+1]\rho)/2}, \sqrt{\log(N)}\}$

(ii) $\| \sum_{k=1}^N \beta_{k+1}^i \mathcal{D}_{k+2}^{(i)*}\|_2 \le b^{\text{D.5}} \max\{N^{(1-i\rho)/2}, \sqrt{\log(N)}\}$

(iii) $\| \sum_{k=1}^N \beta_{k+1}^{m+1} (\mathcal{W}_{k+2}^{\Upsilon^{(m)}} - \mathcal{W}_{k+2}^{\Upsilon^{(m)*}})\|_2 \le b^{\text{D.5}} (N^{(1-[2m+2]\rho)/2} + 1)$

(iv) $\| \sum_{k=1}^N \beta_{k+1}^{m+1} \mathcal{W}_{k+2}^{\Upsilon^{(m)*}}\|_2 \le b^{\text{D.5}} (N^{(1-(m+1)\rho)/2} + 1)$

Proof. Parts (i) and (ii) follow similarly to Lemma C.5: using the martingale difference property, we have from Corollary B.7 (i) that there is a constant $b^{\text{D.5}}$ depending upon Ψ_0 and ρ such that

$$\begin{aligned}
\Big\| \sum_{k=1}^{N} \beta_{k+1}^i (\mathcal{D}_{k+2}^{(i)} - \mathcal{D}_{k+2}^{(i)*}) \Big\|_2^2 &= \sum_{k=1}^{N} |\beta_{k+1}^{2i}| \|\mathcal{D}_{k+2}^{(i)} - \mathcal{D}_{k+2}^{(i)*}\|_2^2 \le \sum_{k=1}^{N} b^{\text{D.5}} |\beta_{k+1}^{2i}| \|\tilde{\theta}_n\|_2^2 \\
&\le \sum_{k=1}^{N} b^{\text{D.5}} \max\{(k+1)^{-(2i+1)\rho}, (k+1)^{-1}\} \\
\Big\| \sum_{k=1}^{N} \beta_{k+1}^i \mathcal{D}_{k+2}^{(i)*} \Big\|_2^2 &= \sum_{k=1}^{N} |\beta_{k+1}^{i}|^2 \|\mathcal{D}_{k+2}^{(i)*}\|_2^2 \le b^{\text{D.5}} \sum_{k=1}^{N} |\beta_{k+1}^{2i}| \\
&\le b^{\text{D.5}} \sum_{k=1}^{N} \max\{(k+1)^{-2i\rho}, (k+1)^{-1}\}
\end{aligned}$$

Thm. 2.2 and Lemma C.3 (ii) establish the final bounds for a potentially larger $b^{\text{D.5}}$. The proof of parts (iii) and (iv) is identical to the proof of (i) and (ii). ☐

The martingales $\{H_{N+2}, H_{N+2}^*\}$ defined in (52) admit attractive bounds:

Corollary D.6. *Under the assumptions of Lemma D.5, the following bounds hold for a constant $b^{\text{D.6}}$ depending upon Ψ_0 and ρ:*

$$\|H_{N+2} - H_{N+2}^*\|_2 \le b^{\text{D.6}} \max\{N^{(1-\rho)/2}, \sqrt{\log(N)}\}$$

☐

E Asymptotic Statistics for Linear Stochastic Approximation

In the special case of linear SA, obtaining a representation for the target bias directly translates to a bias representation for PR averaged estimates because linearization of $\bar{f}$ is no longer needed: $\bar{f}(\theta) = A^*\tilde{\theta}$. The next lemma is a version of Lemma C.1 for PR averaging.

Lemma E.1. *The following holds for the PR averaged estimate* (12),

$$A^* \tilde{\theta}_N^{\text{PR}} = \frac{1}{N}(S_N^{\tau} - S_{N+1}^{\Delta}) \tag{53}$$

where S_N^{τ} and S_N^{Δ} are defined exactly as in Lemma C.1. ☐

E.1 Bias

We begin by analyzing bias within the AD noise setting.

Lemma E.2. *Suppose that the sequence $\{\nu_n\} \subseteq \mathbb{R}^d$ satisfies the recursion:*

$$\nu_{n+1} = (I + \alpha_{n+1}A^*)\nu_n + \alpha_{n+1}\varepsilon_{n+1} \tag{54}$$

where A^* is Hurwitz, $\{\alpha_n\}$ is as defined by (A1) and $\{\varepsilon_{n+1}\}$ is a deterministic sequence. Then, for a constant $b^{\mathrm{E.2}}$ and $n_b > 0$ sufficiently large,

$$\|\nu_{n+1}\|_{\mathsf{M}} \le (1 - b^{\mathrm{E.2}}\tfrac{1}{2}\alpha_{n+1})\|\nu_n\|_{\mathsf{M}} + \|\varepsilon_{n+1}\|_{\mathsf{M}}\,, \quad n \ge n_b$$

in which $\mathsf{M} > 0$ is the unique positive definite matrix solving the Lyapunov equation, $[A^*]^\intercal\mathsf{M} + \mathsf{M}[A^*] = -I$.

Proof. The triangle inequality gives

$$\|\nu_{n+1}\|_{\mathsf{M}} \le \|(I + \alpha_{n+1}A^*)\nu_n\|_{\mathsf{M}} + \alpha_{n+1}\|\varepsilon_{n+1}\|_{\mathsf{M}}$$

Letting λ_1 and λ_d denote the eigenvalues of M with maximal and minimal real part, respectively, we obtain the upper bound: for a constant $b^{\mathrm{E.2}} < \infty$,

$$\begin{aligned}\|(I + \alpha_{n+1}A^*)\nu_n\|_{\mathsf{M}}^2 &= \nu_n^\intercal(I + \alpha_{n+1}A^*)^\intercal\mathsf{M}(I + \alpha_{n+1}A^*)\nu_n \\ &= \|\nu_n\|_{\mathsf{M}}^2 - \alpha_{n+1}\|\beta_n\|_{\mathsf{M}}^2 + \alpha_{n+1}^2\|A^*\beta_n\|_{\mathsf{M}}^2 \\ &\le \|\nu_n\|_{\mathsf{M}}^2\Big(1 - \frac{1}{\lambda_1}\alpha_{n+1} + \alpha_{n+1}^2 b^{\mathrm{E.2}}\frac{\lambda_1}{\lambda_d}\Big)\end{aligned}$$

Since M is positive definite, an application of the bound $\sqrt{1-\delta} \le 1 - \frac{1}{2}\delta$ with $\delta < \frac{1}{\lambda_1}$ completes the proof for $n \ge n_b$ with n_b sufficiently large. ☐

Proof of part (i) of Thm. 2.4. Applying the identity $\bar{f}(\theta) = A^*\tilde{\theta}$ to (6), taking expectations of both sides and rearranging terms gives

$$\mathsf{E}[\tilde{\theta}_{n+1}] = (I + \alpha_{n+1}A^*)\mathsf{E}[\tilde{\theta}_n] + \alpha_{n+1}\mathsf{E}[\Delta_{n+1}]$$

An application of Lemma E.2 with $\nu_n = \mathsf{E}[\tilde{\theta}_n]$ and $\varepsilon_n = \mathsf{E}[\Delta_n]$, yields the contraction: for $n_b > 0$ sufficiently large

$$\mathcal{E}_{n+1} \le (1 - b^{\mathrm{E.2}}\tfrac{1}{2}\alpha_{n+1})\mathcal{E}_n + \alpha_{n+1}\mathcal{E}_n^{\Delta}\,, \quad n \ge n_b$$
$$\text{with } \mathcal{E}_n := \|\nu_n\|_{\mathsf{M}}\,, \quad \mathcal{E}_n^{\Delta} := \|\varepsilon_n\|_{\mathsf{M}}$$

By induction, we obtain the following upper bound: for each $n \ge n_b$,

$$\mathcal{E}_n \le \Xi_{n,n_b}\mathcal{E}_{n_b} + \sum_{k=n_b}^{n} \alpha_k^{n-k}\Xi_{n,k+1}\mathcal{E}_n^{\Delta}\,, \quad \Xi_{n,k} := \prod_{i=k}^{n}(1 - b^{\mathrm{E.2}}\tfrac{1}{2}\alpha_{i+1})$$

Applying the bound $(1+\delta) \le \exp(\delta)$ to the above identity with $\delta < \Xi$ gives, for a constant $\varrho^{2.4}$,

$$\mathcal{E}_n \le \mathcal{E}_{n_b}\exp(-\varrho^{2.4}(\tau_n^b - \tau_{n_b}^b)) + \sum_{k=n_b}^{n} \alpha_k^{n-k}\exp(-\varrho^{2.4}(\tau_n^b - \tau_{k+1}^b))\mathcal{E}_n^{\Delta}$$

in which $\displaystyle\sum_{k=1}^{n} \alpha_k \le \tau_n^b := \alpha_0(1 + (1-\rho)^{-1}[n^{1-\rho} - 1])$.

Since the noise is additive, (13) yields $\mathsf{E}[\Delta_k] = \mathsf{E}[-\mathcal{T}_{k+1} + \mathcal{T}_k]$, in which $\mathcal{T}_k$ is only a function of Φ_k. The drift condition (DV3) implies that $\mathcal{E}_n^{\Delta} \to 0$ geometrically fast (see Thm. A.1 (i)), completing the proof. ☐

When f is affine in θ, the solutions $\hat{f}$ to Poisson's equation (29) with forcing function f takes the form $\hat{f}(\theta, \Phi) = \widehat{A}(\Phi)\theta + \widehat{b}(\Phi)$. Consequently, Υ_k is affine in θ for each k:

$$\Upsilon_{k+2} = -\widehat{A}_{k+2}A_{k+1}\theta_k + \widehat{A}_{k+2}b_{k+1} \tag{55a}$$
$$\Upsilon_{k+2}^* = -\widehat{A}_{k+2}A_{k+1}\theta^* + \widehat{A}_{k+2}b_{k+1} \tag{55b}$$

We note that the definition for Υ^* in (55b) is agrees with the general version in (42). Moreover, the expression in (55b) is the same as the major contributor for bias in linear SA with constant step-size, that is, $\bar{\Upsilon}^*$ in (15) [24].

Proof of part (ii) of Thm. 2.4. Taking expectations of both sides of (53), we obtain $\mathsf{E}[\Delta_{k+1}] = \mathsf{E}[-\mathcal{T}_{k+2} + \mathcal{T}_{k+1} - \alpha_{k+1}\Upsilon_{k+2}]$ from (13), yielding

$$A^*\mathsf{E}[\tilde{\theta}^{\mathsf{PR}}_N] = \frac{1}{N}\Big(\mathsf{E}[S^{\tau}_N] + \mathsf{E}[\mathcal{T}_{N+2} - \mathcal{T}_2] + \sum_{k=1}^{N}\alpha_{k+1}\mathsf{E}[\Upsilon_{k+2}]\Big)$$

The remainder of the proof consists of following the same steps as in the proof of Thm. 2.2 (iv) to obtain the companion to (46):

$$\mathsf{E}[\tilde{\theta}^{\mathsf{PR}}_N] = \alpha_{N+1}\beta_{\theta} + \varepsilon^{\beta}_N\,, \quad \|\varepsilon^{\beta}_N\| \le \begin{cases} b^{\circ}N^{-3\rho/2} & \rho \le 1/2 \\ b^{\circ}N^{\rho/2-1} & \rho > 1/2 \end{cases} \tag{56}$$

with $\beta_{\theta} = (1-\rho)^{-1}[A^*]^{-1}\overline{\Upsilon}^*$. Dividing both sides of the above equation by α_{N+1} and taking limits completes the proof for $\rho \in (0, 2/3)$. ☐

E.2 Asymptotic Covariance

The next lemma will prove itself useful in establishing asymptotic covariances for PR averaging,

Lemma E.3. *Suppose that $\{\mathcal{V}_N\}$ and $\{\mathcal{E}_N\}$ are sequences of random variables such that $\|\mathcal{V}_N\|_2^2 < \infty$ and $\|\mathcal{E}_N\|_2^2 < \infty$. Then, the following approximation holds*

$$\begin{aligned} &\mathsf{Cov}(\mathcal{V}_N + \mathcal{E}_N) = \mathsf{Cov}(\mathcal{V}_N) + \Sigma^{\varepsilon}_N \\ &\textit{where} \quad \Sigma^{\varepsilon}_N \le \varepsilon_N I \quad \textit{with} \quad \varepsilon_N = 2\|\mathcal{V}_N\|_{2,s}\|\mathcal{E}_N\|_{2,s} + \|\mathcal{E}_N\|^2_{2,s} \end{aligned}$$

☐

Proof of part (iii) of Thm. 2.4 for the AD settting. Adding and subtracting $\frac{1}{N}M^*_{N+2}$ to the right side of (53) yields

$$\begin{aligned} \tilde{\theta}^{\mathsf{PR}}_N &= \mathcal{V}_N + \mathcal{E}_N \\ \text{in which } \mathcal{V}_N &= [A^*]^{-1}\frac{1}{N}M^*_{N+2} \\ \mathcal{E}_N &= [A^*]^{-1}\frac{1}{N}(S^{\tau}_N - S^{\Delta}_{N+1} - M^*_{N+2}) \end{aligned} \tag{57a}$$

The expression (39b) gives $S^{\Delta}_{N+1} = M_{N+2} - \mathcal{T}_{N+2} + \mathcal{T}_2$ for the additive noise setting. Lemmas C.4 and C.5 result in the following upper bounds: for a constant b depending upon Ψ_0 and ρ, $\|\mathcal{V}_N\|_2 \le bN^{-1/2}$ and

$$\begin{aligned} \|\mathcal{E}_N\|_2 &\le \|[A^*]^{-1}\|_F\frac{1}{N}(\|S^{\tau}_N\|_2 + \|M_{N+2} - M^*_{N+2}\|_2 + \|\mathcal{T}_{N+2} - \mathcal{T}_2\|_2) \\ &\le b\max\{N^{-(1+\rho)/2}, N^{-(2-\rho)/2}\} \end{aligned} \tag{57b}$$

Consequently, the L_2 norm of $\mathcal{E}_N$ is dominated by $\|S^{\tau}_N\|_2$ for $\rho > 1/2$ and by $\|M_{N+2} - M^*_{N+2}\|_2$ when $\rho < 1/2$.

In view of these bounds and the upper bound after (26b), we take covariances of both sides of (57a) and apply Lemma E.3 to obtain

$$\begin{aligned} \mathsf{Cov}(\tilde{\theta}^{\mathsf{PR}}_N) = \mathsf{Cov}(\mathcal{V}_N + \mathcal{E}_N) &\le \mathsf{Cov}(\mathcal{V}_N) + \varepsilon_N I\,, \\ \varepsilon_N &\le b^{2.4}\max\{N^{-(2+\rho)/2}, N^{-(3-\rho)/2}\} \end{aligned}$$

where $b^{2.4} = 3b^2$. By definition, $M^*_{N+2} = \sum_{i=1}^{N}\mathcal{W}^*_{k+2}$. Applying the martingale difference property, (27) gives

$$\lim_{N\to\infty} N\mathsf{Cov}(\tilde{\theta}^{\mathsf{PR}}_N) = \lim_{N\to\infty} N\mathsf{Cov}(\mathcal{V}_N) = [A^*]^{-1}\Sigma_{\mathcal{W}^*}[A^{*\intercal}]^{-1} \tag{58}$$

☐

In view of Lemma E.1 and (14), the following representation for PR-averaged estimates is obtained:

Proposition E.4. *Under the assumptions of Prop. D.4, the following holds:*

$$A^*\tilde{\theta}_N^{\text{PR}} = -\frac{1}{N}S_N^{\mathcal{A}} - \frac{1}{N}H_{N+2} - \frac{1}{N}J_{N+2} + \frac{1}{N}S_N^u \tag{59}$$

where $\{S_N^u\}$ is a deterministic sequence,

(i) *$\{H_{N+2}\}$ is a martingale satisfying* $\lim_{N\to\infty}\frac{1}{N}\mathsf{Cov}(H_{N+2}) = \Sigma_{\mathcal{W}^*}$

(ii) *For a constant $b^{\text{E.4}}$ depending upon Ψ_0 and ρ, and $N \ge 1$,*

$$\|S_N^{\mathcal{A}}\|_{2,s} \le b^{\text{E.4}}\big[1 + N^{1-\rho}\|\tilde{\theta}_N^{\text{PR}}\|_{2,s} + \sum_{k=1}^{N-1} k^{-\rho}\|\tilde{\theta}_k^{\text{PR}}\|_{2,s}\big]$$

(iii) *If in addition, $m = m(\rho)$ in (14) is chosen so that $m\rho > 1$, then for $N \ge 1$, $\|J_{N+2}\|_2 \le b^{\text{E.4}}N^{\rho/2}$*

Proof. Summing (14) from $k = 1$ to N and substituting it into (53) gives (59) in which

$$S_N^{\mathcal{A}} = \sum_{k=1}^{N}\alpha_{k+1}\bar{G}_{k+1}\tilde{\theta}_k\,, \qquad S_N^u = \sum_{k=1}^{N}\alpha_{k+1}\bar{u}_{k+1}$$

$$J_{N+2} = -\mathcal{T}_{N+2}^\bullet + \mathcal{T}_2^\bullet - S_N^\tau + \sum_{k=1}^{N}[(\mathcal{R}_{k+1}^\bullet - \mathcal{T}_{k+1}^\bullet) + (-\alpha_{k+1})^{m+1}\Upsilon_{k+2}^\bullet]$$

We proceed with the proof of part (i). We have $\mathcal{W}_{k+2}^{(0)*} = \mathcal{W}_{k+2}^*$ from Prop. D.4, which along with (14) and the definition of $\{H_k\}$ before Lemma D.5 gives the identity $\frac{1}{N}H_{N+2} = \mathcal{V}_N + \mathcal{E}_N$, with $\mathcal{V}_N = \frac{1}{N}M_{N+2}^*$ and $\mathcal{E}_N = \mathcal{E}_N^1 + \mathcal{E}_N^2$ where

$$\mathcal{E}_N^1 = \frac{1}{N}(H_{N+2} - H_{N+2}^*)$$

$$\mathcal{E}_N^2 = \frac{1}{N}\sum_{k=1}^{N}\Big[\sum_{i=1}^{m}\mathcal{D}_{k+2}^{(i)*} - \alpha_{k+1}\mathcal{W}_{k+2}^{\Upsilon^{(0)*}} - \beta_{k+1}^{m+1}\mathcal{W}_{k+2}^{\Upsilon^{(m)*}}\Big]$$

Moment bounds on $\mathcal{V}_N$ are obtained exactly as it was done in the proof of Thm. 2.4 (iii) for the AD setting, while bounds on $\mathcal{E}_N^1$ follow from Prop. E.4 (i). They are of the form $\|\mathcal{V}_N\|_2 \le bN^{-1/2}$ and $\|\mathcal{E}_N^1\|_2 \le b\max\{N^{-(1+\rho)/2}, \sqrt{\log(N)}/N\}$, in which b is a constant depending upon Ψ_0 and ρ.

Moreover, similar arguments as the ones in Lemma D.5 (ii) and (iv) yield the following, for a potentially larger b,

$$\Big\|\sum_{k=1}^{N}\sum_{i=1}^{m}\mathcal{D}_{k+2}^{(i)*}\Big\|_2 \le b\max\{N^{(1-\rho)/2}, \sqrt{\log(N)}\}$$

$$\Big\|\beta_{k+1}^{m+1}\mathcal{W}_{k+2}^{\Upsilon^{(m)*}}\Big\|_2 \le b(N^{(1-(m+1)\rho)/2} + 1)$$

$$\Big\|\sum_{k=1}^{N}\alpha_{k+1}\mathcal{W}_{k+2}^{\Upsilon^{(0)*}}\Big\|_2 \le b\max\{N^{(1-\rho)/2}, \sqrt{\log(N)}\}$$

where the first bound holds since the term associated with $i = 1$ dominates. This implies $\|\mathcal{E}_N^2\|_2 \le 3b\max\{N^{-(1+\rho)/2}, \sqrt{\log(N)}/N\}$.

Then, Lemma E.3 and the upper bound after (26b) yield

$$\mathsf{Cov}\Big(\frac{1}{N}H_{N+2}\Big) = \mathsf{Cov}\Big(\mathcal{V}_N + \mathcal{E}_N\Big) \le \mathsf{Cov}(\mathcal{V}_N) + \varepsilon_N I\,,$$
$$\varepsilon_N \le b^{2.4}\max\{N^{-1-\rho}, \log(N)/N^2\}$$

where $b^{2.4} = 15b^2$. Multiplying both sides of the above equation by N and taking limits completes the proof for part (i):

$$\lim_{N\to\infty}\frac{1}{N}\mathsf{Cov}(H_{N+2}) = \lim_{N\to\infty}N\mathsf{Cov}(\mathcal{V}_N) = \Sigma_{\mathcal{W}^*}$$

We now turn to the proof of part (ii). Writing $\tilde{\theta}_k = k\tilde{\theta}^{\sf PR}_k - (k-1)\tilde{\theta}^{\sf PR}_{k-1}$ gives

$$
\begin{aligned}
S^{\mathcal{A}}_N &= \sum_{k=1}^{N} \alpha_{k+1}\bar{G}_{k+1}[k\tilde{\theta}^{\sf PR}_k - (k-1)\tilde{\theta}^{\sf PR}_{k-1}] \\
&= \alpha_{N+1}\bar{G}_{N+1}N\tilde{\theta}^{\sf PR}_N - \alpha_2\bar{G}_2\tilde{\theta}^{\sf PR}_1 - \sum_{k=1}^{N-1}(\alpha_{k+1}\bar{G}_{k+1} - \alpha_k\bar{G}_k)k\tilde{\theta}^{\sf PR}_k
\end{aligned}
$$

where the last equality was obtained through Lemma C.2.

Taking L_2 span norms of both sides of the above equation and using the triangle inequality

$$
\begin{aligned}
\|S^{\mathcal{A}}_N\|_{2,s} \le |\alpha_{N+1}|\|\bar{G}_{N+1}\|_F N\|\tilde{\theta}^{\sf PR}_N\|_{2,s} + |\alpha_2|\|\bar{G}_2\|_F\|\tilde{\theta}^{\sf PR}_1\|_{2,s} \\
+ \sum_{k=1}^{N-1}\|\alpha_{k+1}\bar{G}_{k+1} - \alpha_k\bar{G}_k\|_F k\|\tilde{\theta}^{\sf PR}_k\|_{2,s}
\end{aligned}
$$

This and Prop. D.4 (iv) completes the proof of (ii).

Part (iii) follows from the definitions of $\{H_{N+2}, J_{N+2}\}$ along with the bounds in Prop. D.4 and Lemma C.4 and corollary D.6.

□

The representation (59) in Prop. E.4 is equivalent to (7), in which $\mathcal{Z}_n := -\sqrt{N}[A^*]^{-1}(H_{N+2} + J_{N+2} + S^{\mathcal{A}}_N)$ and $\beta_N := (N\alpha_{N+1})^{-1}S^u_N$.

The representation (59) is particularly useful in the setting of linear SA since it results in tighter bounds for the L_2 span norm of $\theta^{\sf PR}_N$ when compared to Thm. 2.2 (iii). It will be clear that the asymptotic covariance of the right hand side of (59) is dominated by $\{H_{N+2}\}$ and equals $\Sigma_{\mathcal{W}^*}$.

Bounds on $\|\tilde{\theta}^{\sf PR}_N\|_{2,s}$ are obtained next for the general nonlinear SA algorithm.

Lemma E.5. *Suppose that the assumptions of Prop. D.4 hold. Then,* $\|\tilde{\theta}^{\sf PR}_N\|_{2,s} \le b^{\text{E.5}} N^{-1/2}$*,* $N \ge 1$*, for a constant* $b^{\text{E.5}}$ *depending upon* Ψ_0 *and* ρ*.*

Proof. Part (iii) of Thm. 2.2 and the inequality after (26b) give $\|\tilde{\theta}^{\sf PR}_N\|_{2,s} \le b^{2.2}\max\{\alpha_N, N^{-1/2}\}$, which completes the proof for $\rho \ge 1/2$. We proceed with the proof of the desired bound for $\rho < 1/2$.

For each $\rho \in (0, 1/2)$, we choose m in (14) so that $m\rho > 1$. Then, taking L_2 span norms of both sides of (59) yields the recursive sequence of bounds,

$$
\begin{aligned}
\|A^*\tilde{\theta}^{\sf PR}_N\|_{2,s} &\le \frac{1}{N}\|S^{\mathcal{A}}_N\|_{2,s} + \frac{1}{N}\|H_{N+2}\|_{2,s} + \frac{1}{N}\|J_{N+2}\|_{2,s} \\
&\le b^{\text{E.4}}\frac{1}{N}\Big(N^{1-\rho}\|\tilde{\theta}^{\sf PR}_N\|_{2,s} + \sum_{k=1}^{N-1} k^{-\rho}\|\tilde{\theta}^{\sf PR}_k\|_{2,s}\Big) + b^{\text{E.4}}N^{-1/2}
\end{aligned}
\tag{60}
$$

where the last inequality follows from Prop. E.4, along with $\|S^u_N\|_{2,s} = 0$.

Applying Thm. 2.2 (iii) (the bound (8)) gives $\|\tilde{\theta}^{\sf PR}_k\|^2_{2,s} \le {\sf E}[\|\theta^{\sf PR}_k - \theta^*\|^2] \le b^{2.2}\alpha_k^2$ for $\rho < 1/2$. This bound together with Lemma C.3 (ii) gives

$$
\|A^*\tilde{\theta}^{\sf PR}_N\|_{2,s} \le b^{\text{E.4}}\Big(\frac{1}{1-2\rho} + 1\Big)\alpha_N^2 + b^{\text{E.4}}N^{-1/2}\,, \quad \rho < 1/2 \tag{61}
$$

completing the proof for $\rho \in [1/4, 1/2)$.

If $\rho < 1/4$ then (61) gives $\|\tilde{\theta}^{\sf PR}_k\|_{2,s} \le b^{\text{E.5}}\alpha_k^2$, which is an improvement on Thm. 2.2 (iii). Substituting this into (60) and repeating the above process yields a sequence of recursive bounds similar to (61):

$$
\|A^*\tilde{\theta}^{\sf PR}_N\|_{2,s} \le b^{\text{E.5}}\Big(\frac{1}{1-3\rho} + 1\Big)\alpha_N^3 + b^{\text{E.4}}N^{-1/2}\,, \quad \rho < 1/4
$$

which completes the proof for $\rho \in [1/6, 1/4)$.

Continuing to repeat this process establishes the desired bound for all $\rho \in (0, 1)$. □

Now, we can apply the bounds in Lemma E.5 to Prop. E.4 (ii) and conclude that the sequences $\{S^{\mathcal{A}}_N, J_{N+2}\}$ do not dominate the asymptotic covariance of the right side of (59).

Proof of part (iii) of Thm. 2.4 for the MU settting. It follows from (59) that $\tilde{\theta}_N^{\mathsf{PR}} = \mathcal{V}_N + \mathcal{E}_N$, in which

$$\mathcal{V}_N = -[A^*]^{-1}\frac{1}{N}H_{N+2}\,, \qquad \mathcal{E}_N = -[A^*]^{-1}\frac{1}{N}(J_{N+2} - S_N^u + S_N^{\mathcal{A}}) \tag{62}$$

Applying the bound $\|\tilde{\theta}_N^{\mathsf{PR}}\|_{2,s} \le b^{\mathrm{E.5}} N^{-1/2}$ from Lemma E.5 to Prop. E.4 (ii), we obtain

$$\|S_N^{\mathcal{A}}\|_{2,s} \le b^{\mathrm{E.4}} b^{\mathrm{E.5}} N^{(1-2\rho)/2}$$

Applying Prop. E.4 gives, for a constant b depending upon Ψ_0 and ρ,

$$\begin{aligned}
\|\mathcal{V}_N\|_{2,s} &\le bN^{-1/2} \\
\|\mathcal{E}_N\|_{2,s} &\le \|[A^*]^{-1}\|_F \frac{1}{N}\Big(\|J_{N+2}\|_{2,s} + \|S_N^{\mathcal{A}}\|_{2,s}\Big) \\
&\le b\max\{N^{-1+\rho/2}, N^{-1/2-\rho}\}
\end{aligned}$$

Applying Lemma E.3 we obtain

$$\mathsf{Cov}(\tilde{\theta}_N^{\mathsf{PR}}) = \tfrac{1}{N}\Sigma^{\mathsf{PR}} + \mathcal{E}_N^{\Sigma}\,, \quad \mathrm{tr}\,(\mathcal{E}_N^{\Sigma}) \le \begin{cases} b^{2.4} N^{-[1+\rho]} & \rho \le 1/3 \\ b^{2.4} N^{-[3-\rho]/2} & \rho > 1/3 \end{cases} \tag{63}$$

where $b^{2.4} = 3b^2$. Multiplying both sides of the above equation by N and taking limits completes the proof:

$$\lim_{N\to\infty} N\mathsf{Cov}(\tilde{\theta}_N^{\mathsf{PR}}) = \lim_{N\to\infty} N\mathsf{Cov}(\mathcal{V}_N) = [A^*]^{-1}\Sigma_{\mathcal{W}^*}[A^{*\intercal}]^{-1}$$

□

Proof of Thm. 2.5. Part (i) follows from taking norms of both sides of (56) and applying the triangle inequality.

The proof of part (ii) begins by taking square roots of both sides of (26b) to obtain, via the triangle inequality

$$\|\tilde{\theta}_N^{\mathsf{PR}}\|_2 \le \sqrt{\mathrm{tr}\,(\mathsf{Cov}(\tilde{\theta}_N^{\mathsf{PR}}))} + \|\mathsf{E}[\tilde{\theta}_N^{\mathsf{PR}}]\|$$

Substituting eqs. (56) and (63) into the above equation and using the triangle inequality once more completes the proof. □

F Numerical Experiments

Here we survey additional results and provide further details on the numerical experiments discussed in Section 3.

Similarly to what is shown by Fig. 1, Figs. 3 and 4 display the empirical mean and variance for the final PR-averaged estimates $\theta_N^{\mathsf{PR}\,i}$ across all independent runs as functions of ρ for $a \in \{0.3, 0.5\}$. Additionally, ratios between these empirical statistics and their associated theoretical asymptotic predictions are also shown by Fig. 5 for each a. Note that there is no comparison with theoretical values for the empirical mean when $a = 0.5$, in which case the Markov chain is i.i.d. and bias vanishes very fast, by Thm. 2.4 (i).

The results in Figs. 3 and 4 agree with the discussion surrounding Fig. 1: poor solidarity with theory is observed in both large and small-ρ regimes irrespective of the choice of a. Again, this is justified by the error bounds in Thm. 2.5.

Impact of spectral gap. The transition matrix P has eigenvalues $\{\lambda_1, \lambda_2\} = \{1, 2a-1\}$, and the spectral gap is thus $\gamma = 1 - |\lambda_2| = 1 - |2a-1|$. It is maximized when $a = 0.5$ and vanishes if either $a \sim 0$ or $a \sim 1$. It is commonly assumed that a small spectral gap is a sign of difficulty in numerical algorithms, yet a glance at Lemma 3.1 shows that:

(i) While it is true that $\bar{\Upsilon}^*$ increases without bound when $a \uparrow 1$, when $a \downarrow 0$ the spectral gap vanishes yet $\bar{\Upsilon}^*$ converges to a finite value.

(ii) The asymptotic covariance *vanishes* as $a \downarrow 0$, even as the spectral gap vanishes. In conclusion: long statistical memory is not always detrimental to algorithmic performance.

The above conclusion is in agreement to what can be observed by the results in Fig. 5 since the empirical covariance for $a = 0.3$ is closer to optimal than for the i.i.d. case ($a = 0.5$). Moreover, the setting $a = 0.3$ also outperforms $a = 0.7$ in terms of empirical variance even though these two choices result in the same spectral gap.

When it comes to bias, it appears that $a = 0.3$ and $a = 0.5$ have similar performances: based upon Thm. 2.4 (i), bias for the i.i.d. case vanishes very fast, so the low magnitude observed for the empirical mean with $a = 0.5$ is expected. The curve associated with $a = 0.3$ seems to track the prediction in Thm. 2.4 (ii) well only facing problems in the large ρ regime. Tracking for case $a = 0.7$ is not as good as when $a = 0.3$ but also appears to be consisted of with theory. However, tracking is poor in both the low and large ρ regimes.

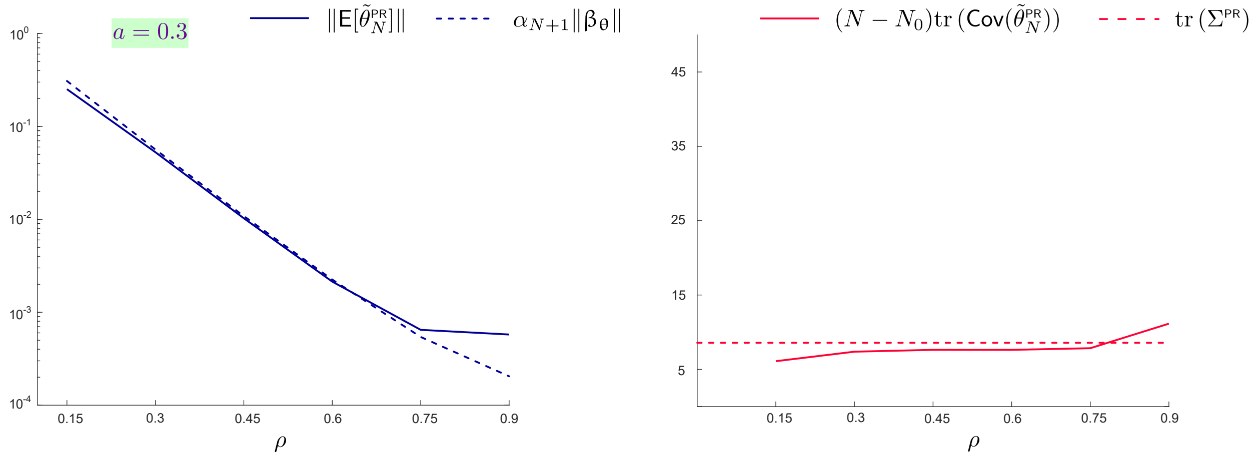

Figure 3: Empirical and theoretical mean and variance with $a = 0.3$.

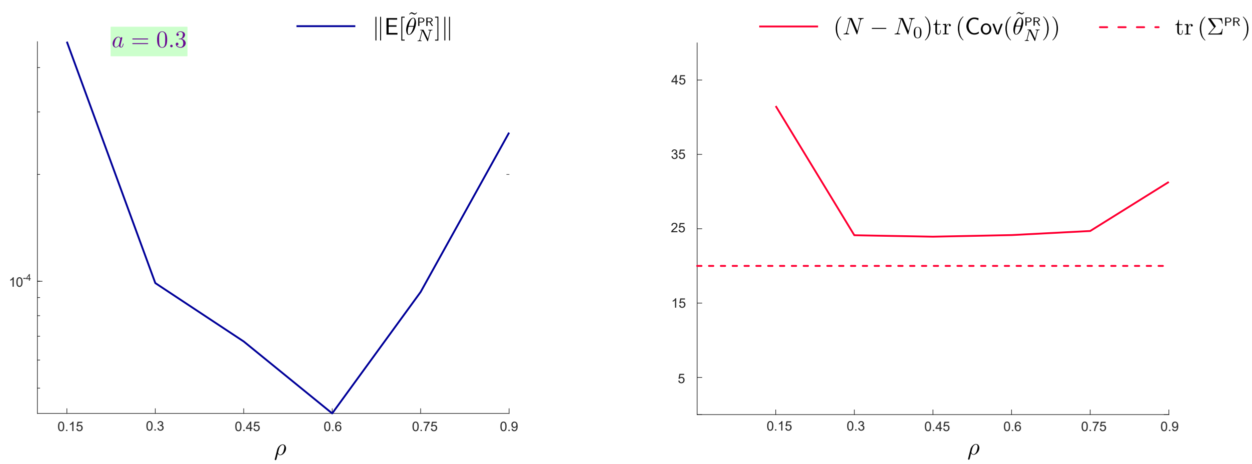

Figure 4: Empirical and theoretical mean and variance with $a = 0.5$.

Identifying the constants in (17b). The example considered in Section 3 is of the form (16) of [8], in which $X_k = \Phi_k$ and $\mathcal{W}_{k+1}^0 \equiv 0$. This prior work requires a quadratic Lyapunov function for the mean-flow vector field (see discussion in Section 1.1). The values in (24) were selected for the model (22) so that $V(\theta) = \frac{1}{2}\|\theta\|^2$ satisfies (17a) with $c_0 = 1$. The final term in (17b) is thus

$$L = 520\alpha_0(\|\theta^*\| + 1)^2$$

We conclude this section with a proof of Lemma 3.1, followed by details on the computing resources used to perform the experiments in Section 3.

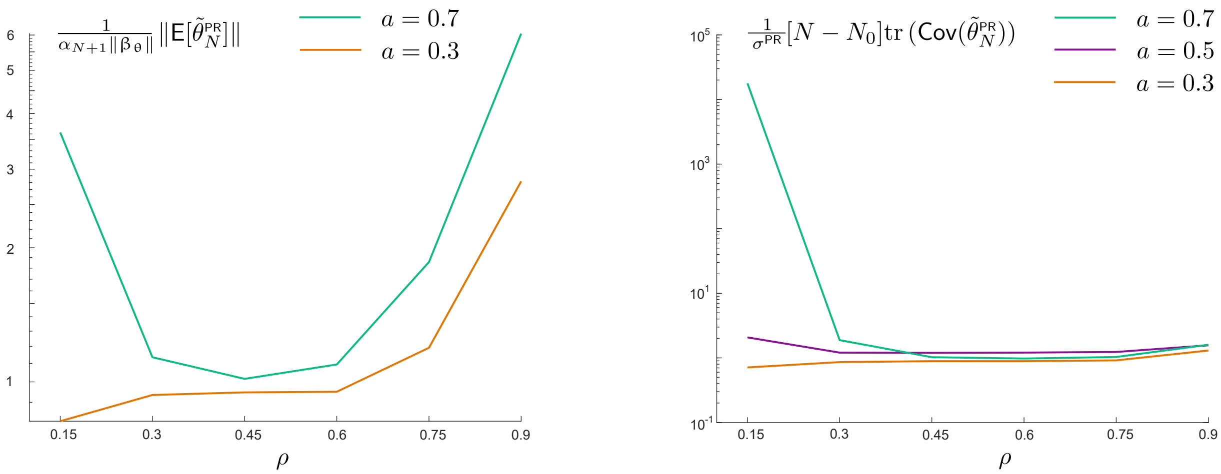

Figure 5: Ratios between empirical and theoretical mean and variance for different choices of a.

Proof of Lemma 3.1. Let $\hat{g} : \mathsf{X} \to \mathbb{R}$ denote the solution to Poisson's equation (25) with forcing function $g(x) \equiv x$. Since $\uppi$ is uniform, Poisson's equation equation gives for each $x \in \mathsf{X}$,

$$\mathsf{E}[\hat{g}(\Phi_{k+1})|\Phi_k = x] = a\hat{g}(x) + (1-a)\hat{g}(\overline{x}) = \hat{g}(x) - x + \tfrac{1}{2}$$

where $\overline{x} = \mathsf{X}\setminus\{x\}$. Solutions to (29) are unique up to an additive constant, so let $\hat{g}(1) = 1$. Then, substituting $x = 1$ into the above equation gives $\hat{g}(0) = \frac{1}{2}\frac{1-2a}{1-a}$. It follows that $f(\theta, x)$ can be re-written as

$$f(\theta, x) = [x(A^1 - A^0) + A^0]\theta - [x(b^1 - b^0) + b^0] \tag{64a}$$

which gives

$$\hat{f}(\theta, x) = \hat{g}(x)[(A^1 - A^0)\theta - (b^1 - b^0)] \tag{64b}$$

$$\partial_\theta \hat{f}(\theta, x) = \hat{g}(x)(A^1 - A^0) \tag{64c}$$

where again we have used the fact that solutions to Poisson's equation are unique up to an additive constant.

The definition of Υ^* in (42) implies

$$\begin{aligned}
\Upsilon^*_{k+2} &= -\hat{g}(\Phi_{k+2})(A^1 - A^0)[A_{k+1}\theta^* - b_{k+1}] \\
\bar{\Upsilon}^* = \mathsf{E}[\Upsilon^*_{k+2}] &= -\tfrac{1}{2}(A^1 - A^0)\Big[(A^0\theta^* - b^0)\mathsf{E}[\hat{g}(\Phi_{k+2})|\Phi_{k+1} = 0] + (A^1\theta^* - b^1)\mathsf{E}[\hat{g}(\Phi_{k+2})|\Phi_{k+1} = 1]\Big] \\
&= -\tfrac{1}{2}(A^1 - A^0)\Big[(A^0\theta^* - b^0)\Big(\hat{g}(0) - g(0) + \tfrac{1}{2}\Big) + (A^1\theta^* - b^1)\Big(\hat{g}(1) - g(1) + \tfrac{1}{2}\Big)\Big]
\end{aligned}$$

Part (i) is established upon using this with the following identity, obtained from (22b):

$$0 = \bar{f}(\theta^*) = \tfrac{1}{2}[(A^0 + A^1)\theta^* - b^0 - b^1] \tag{65}$$

For part (ii), it follows from (28) that we can write the asymptotic covariance Σ_Δ as

$$\Sigma_\Delta = \mathsf{E}_\pi[\Delta^*_k \hat{\Delta}^{*\intercal}_k + \hat{\Delta}^*_k \Delta^{*\intercal}_k - \Delta^*_k \Delta^{*\intercal}_k]$$

where $\Delta^*_k = f(\theta^*, \Phi_k)$ and $\hat{\Delta}^*_k = \hat{f}(\theta^*, \Phi_k)$. The law of total expectation and the definitions in (64) give

$$\begin{aligned}
\mathsf{E}_\pi[\hat{\Delta}^*_k \Delta^{*\intercal}_k] &= \tfrac{1}{2}(\mathsf{E}_\pi[\hat{\Delta}^*_k \Delta^{*\intercal}_k|\Phi_k = 0] + \mathsf{E}_\pi[\hat{\Delta}^*_k \Delta^{*\intercal}_k|\Phi_k = 1]) \\
&= \tfrac{1}{2}[(A^1 - A^0)\theta^* - (b^1 - b^0)][\hat{g}(0)(A^0\theta^* - b^0)^\intercal + \hat{g}(1)(A^1\theta^* - b^1)^\intercal] \\
\mathsf{E}_\pi[\Delta^*_k \Delta^{*\intercal}_k] &= \tfrac{1}{2}(\mathsf{E}_\pi[\Delta^*_k \Delta^{*\intercal}_k|\Phi_k = 0] + \mathsf{E}_\pi[\Delta^*_k \Delta^{*\intercal}_k|\Phi_k = 1]) \\
&= \tfrac{1}{2}[(A^0\theta^* - b^0)(A^0\theta^* - b^0)^\intercal + (A^1\theta^* - b^1)(A^1\theta^* - b^1)^\intercal]
\end{aligned}$$

Again, using (65) we obtain

$$\mathsf{E}_\pi[\hat{\Delta}^*_k \Delta^{*\intercal}_k] = -(A^0\theta^*)(A^0\theta^*)^\intercal(\hat{g}(0) - \hat{g}(1))\,, \quad \mathsf{E}_\pi[\Delta^*_k \Delta^{*\intercal}_k] = (A^0\theta^*)(A^0\theta^*)^\intercal$$

which completes the proof of (ii). $\square$