Prelimit Coupling and Steady-State Convergence of Constant-stepsize Nonsmooth Contractive SA

Yixuan Zhang
University of Wisconsin-Madison
Madison, Wisconsin, USA
yzhang2554@wisc.edu

Dongyan (Lucy) Huo
Cornell University
Ithaca, New York, USA
dh622@cornell.edu

Yudong Chen
University of Wisconsin-Madison
Madison, Wisconsin, USA
yudong.chen@wisc.edu

Qiaomin Xie
University of Wisconsin-Madison
Madison, Wisconsin, USA
qiaomin.xie@wisc.edu

ABSTRACT

Motivated by Q-learning, we study nonsmooth contractive stochastic approximation (SA) with constant stepsize. We focus on two important classes of dynamics: 1) nonsmooth contractive SA with additive noise, and 2) synchronous and asynchronous Q-learning, which features both additive and multiplicative noise. For both dynamics, we establish weak convergence of the iterates to a stationary limit distribution in Wasserstein distance. Furthermore, we propose a prelimit coupling technique for establishing steady-state convergence and characterize the limit of the stationary distribution as the stepsize goes to zero. Using this result, we derive that the asymptotic bias of nonsmooth SA is proportional to the square root of the stepsize, which stands in sharp contrast to smooth SA. This bias characterization allows for the use of Richardson-Romberg extrapolation for bias reduction in nonsmooth SA.[1]

ACM Reference Format:

Yixuan Zhang, Dongyan (Lucy) Huo, Yudong Chen, and Qiaomin Xie. 2024. Prelimit Coupling and Steady-State Convergence of Constant-stepsize Nonsmooth Contractive SA. In *Abstracts of the 2024 ACM SIGMETRICS/IFIP PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS/PERFORMANCE Abstracts '24), June 10–14, 2024, Venice, Italy*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3652963.3655076

1 INTRODUCTION

Stochastic Approximation (SA) is a fundamental algorithmic paradigm for solving fixed-point problems iteratively based on noisy observations. A typical SA algorithm is of the form

$$\theta_{t+1}^{(\alpha)} = \theta_t^{(\alpha)} + \alpha\big(\widetilde{\mathcal{H}}(\theta_t^{(\alpha)}, w_t) - \theta_t^{(\alpha)}\big), \tag{1}$$

where $\{w_t\}_{t\geq 0}$ represent the i.i.d. noise sequence and $\alpha > 0$ is a constant stepsize. The SA procedure (1) aims to approximately find the solution θ^* to the fixed-point equation $\mathcal{H}(\theta^*) = \theta^*$, where $\mathcal{H}(\cdot) := \mathbb{E}_w[\widetilde{\mathcal{H}}(\cdot, w)]$ is the expectation of the operator $\widetilde{\mathcal{H}}(\cdot, w)$ with respect to the noise. Equation (1) covers many popular algorithms, such as the prevalent stochastic gradient descent (SGD) algorithm for minimizing an objective function [6], and variants of TD-learning algorithms for policy evaluation in reinforcement learning (RL).

In this work, we focus on *nonsmooth* contractive SA, where the operator $\widetilde{\mathcal{H}}(\cdot, w)$ may be nondifferentiable (in its first argument) and $\mathcal{H}(\cdot)$ is a contractive mapping with respect to a norm $\|\cdot\|_c$. One prominent example of nonsmooth contractive SA is the celebrated Q-learning algorithm for optimal control in RL [10], where $\widetilde{\mathcal{H}}$ corresponds to the noisy optimal Bellman operator involving a max function. It is of fundamental interest to gain a complete understanding of the evolution and long-run behavior of the iterates $\{\theta_t^{(\alpha)}\}_{t\geq 0}$ generated by nonsmooth contractive SA.

Under suitable conditions on the operator $\widetilde{\mathcal{H}}$ and the noise sequence $\{w_t\}_{t\geq 0}$, the SA iterates $\{\theta_t^{(\alpha)}\}_{t\geq 0}$ form a time-homogeneous Markov chain and quickly converge to some limit random variable $\theta^{(\alpha)}$ [1, 11]. Recent work has developed a suite of results for *smooth* SA [1, 2, 4], including the geometric convergence of the chain, finite-time bounds on the higher moments, as well as properties of the limit $\theta^{(\alpha)}$. It has been observed that often $\mathbb{E}[\theta^{(\alpha)}] \neq \theta^*$, due to the use of constant stepsize. The difference $\mathbb{E}[\theta^{(\alpha)}] - \theta^*$ is referred to as the asymptotic bias. In particular, for SA with *differentiable* dynamic, the work [1, 4] makes use of Taylor expansion of $\widetilde{\mathcal{H}}$ to establish that the asymptotic bias is proportional to the stepsize α (up to a higher order term), i.e.,

$$\mathbb{E}[\theta^{(\alpha)}] - \theta^* = c\alpha + o(\alpha), \tag{2}$$

where c is some vector independent of α and $o(\alpha)$ denotes a term that decays faster than α. Such a fine-grained characterization of SA iterates gives rise to variance and bias reduction techniques that lead to improved estimation of the target solution θ^*, as well as efficient statistical inference procedures [1, 4, 5].

For nonsmooth SA, far little is known. Existing analysis based on the linearization / Taylor expansion of $\widetilde{\mathcal{H}}$ is no longer applicable. Hence, distributional convergence and bias characterization results like (2) have not been established for nonsmooth SA procedures like Q learning. In fact, it is not even clear whether equation (2) remains valid for nonsmooth SA, and if not, what is the correct characterization.

[1] Extended Abstract. The full paper can be found at [12].

SIGMETRICS/PERFORMANCE Abstracts '24, June 10–14, 2024, Venice, Italy.

ACM ISBN 979-8-4007-0624-0/24/06.
https://doi.org/10.1145/3652963.3655076

2 MAIN RESULTS

To investigate the above questions, we consider two important classes of nonsmooth contractive SA algorithms:

(i) Nonsmooth SA with additive noise, where $\widetilde{\mathcal{H}}(\theta, w) \equiv \mathcal{T}(\theta)+w$. Our results cover operators $\mathcal{T}$ that are $g \circ F$ decomposable, which is a rich class of smooth and nonsmooth functions [9].

(ii) A general form of Q-learning, which are nonsmooth SA with both additive and multiplicative noise. The model covers both synchronous Q-learning and asynchronous Q- learning.

The first main result establishes the weak convergence of the Markov chain $\{\theta_t^{(\alpha)}\}_{t \geq 0}$ in W_2 — the Wasserstein distance of order 2 with respect to the contraction norm $\|\cdot\|_c$.

Theorem 1 (Informal). *$\{\theta_t^{(\alpha)}\}_{t \geq 0}$ converges to a unique stationary distribution $\theta^{(\alpha)}$ in W_2 for both settings (i) and (ii).*

Moreover, we characterize the geometric convergence rate. As a by-product of our analysis, we derive finite-time upper bounds on $\mathbb{E}\|\theta_t^{(\alpha)} - \theta^*\|_c^{2n}$, the $2n$-th moments of the estimation errors.

We next turn to the characterization of the stationary distribution of $\{\theta_t^{(\alpha)}\}_{t \geq 0}$. Existing techniques, which are based on linearizing $\widetilde{\mathcal{H}}(\theta, w)$ as $\theta \to \theta^*$, are not applicable for nonsmooth SA. We take an alternative approach by studying the limiting behavior of the diffusion-scaled iterates $Y_t^{(\alpha)} := \frac{\theta_t^{(\alpha)} - \theta^*}{\sqrt{\alpha}}$ as the constant stepsize α approaches 0. The weak convergence of $\theta_t^{(\alpha)}$ to a limit $\theta^{(\alpha)}$ implies that $Y_t^{(\alpha)}$ converges weakly to the limit $Y^{(\alpha)} := \frac{\theta^{(\alpha)} - \theta^*}{\sqrt{\alpha}}$ as $t \to \infty$. Therefore, to understand the stationary distribution $\theta^{(\alpha)}$ and its scaled version $Y^{(\alpha)}$, we are interested in characterizing *steady-state convergence*, i.e., the convergence of $Y^{(\alpha)}$ as $\alpha \to 0$ and the limit Y (if exists). This limit is illustrated by the red solid path in Fig. 1.

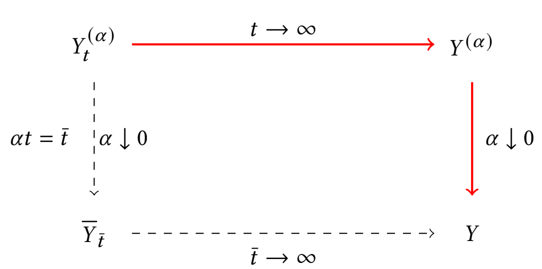

Figure 1: Steady-state convergence.

Existing approaches to steady-state convergence face severe challenges in the nonsmooth SA setting. In this work, we develop a new *prelimit coupling technique*, which allows us to establish the desired result.

Theorem 2 (Informal). *$Y^{(\alpha)}$ converges in W_2 to a unique limiting random variable Y as $\alpha \to 0$*

Importantly, our technique can handle both additive noise and multiplicative noise, and provide an explicit rate of convergence. We remark that our technique can be potentially applied to the study of steady-state convergence in other stochastic dynamical systems and hence may be of its own interest.

The convergence of $Y^{(\alpha)}$ in W_2 implies convergence of $\mathbb{E}[Y^{(\alpha)}]$ and hence that of $\mathbb{E}[\theta^{(\alpha)}]$. Consequently, we can characterize the asymptotic bias of the SA iterates $\theta_t^{(\alpha)}$ and further relate fine-grained properties of the bias to the structure of the SA update (1).

Theorem 3 (Informal). *The asymptotic bias satisfies*

$$\mathbb{E}[\theta^{(\alpha)}] - \theta^* = \mathbb{E}[Y] \cdot \sqrt{\alpha} + o(\sqrt{\alpha}). \tag{3}$$

Moreover, $\mathbb{E}[Y] \neq 0$ when the operator $\widetilde{\mathcal{H}}$ is truly nonsmooth.

Therefore, the asymptotic bias is of order $\sqrt{\alpha}$ precisely when the SA update is nonsmooth. This result stands in sharp contrast to the α-order bias of smooth SA in equation (2).

Finally, we explore the implications of the above results for iterate averaging and extrapolation. In particular, we apply Polyak-Ruppert (PR) tail averaging [7, 8] and Richardson-Romberg (RR) extrapolation [3] to the iterates generated by contractive SA algorithms. We establish the following guarantees on the resulting estimation errors and biases in the presence of nonsmoothness.

Theorem 4 (Informal). *Thanks to the bias characterization in (3), the RR extrapolation technique can be employed to eliminate the leading term $\mathbb{E}[Y] \cdot \sqrt{\alpha}$ and reduce the asymptotic bias $o(\sqrt{\alpha})$.*

ACKNOWLEDGMENTS

We acknowledge support from NSF grants CCF-2047910, CCF-1704828, CNS-1955997, EPCN-2339794 and DMS-2023239.

REFERENCES

[1] Aymeric Dieuleveut, Alain Durmus, and Francis Bach. 2020. Bridging the gap between constant step size stochastic gradient descent and Markov chains. *The Annals of Statistics* 48, 3 (2020), 1348 – 1382. https://doi.org/10.1214/19-AOS1850
[2] Alain Durmus, Pablo Jiménez, Eric Moulines, and Salem Said. 2021. On Riemannian Stochastic Approximation Schemes with Fixed Step-Size. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*. PMLR, 1018–1026. https://proceedings.mlr.press/v130/durmus21a.html
[3] F B Hildebrand. 1987. *Introduction to numerical analysis* (2 ed.). Dover Publications, Mineola, NY.
[4] Dongyan (Lucy) Huo, Yudong Chen, and Qiaomin Xie. 2023. Bias and extrapolation in Markovian linear stochastic approximation with constant stepsizes. In *Abstract Proceedings of the 2023 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*. 81–82. https://doi.org/10.1145/3578338.3593526
[5] Dongyan Lucy Huo, Yudong Chen, and Qiaomin Xie. 2024. Effectiveness of Constant Stepsize in Markovian LSA and Statistical Inference. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38. 20447–20455.
[6] Guanghui Lan. 2020. *First-order and stochastic optimization methods for machine learning*. Vol. 1. Springer.
[7] Boris T. Polyak and Anatoli B. Juditsky. 1992. Acceleration of Stochastic Approximation by Averaging. *SIAM Journal on Control and Optimization* 30, 4 (1992), 838–855. https://doi.org/10.1137/0330046
[8] David Ruppert. 1988. *Efficient estimations from a slowly convergent Robbins-Monro process*. Technical Report. Cornell University Operations Research and Industrial Engineering.
[9] Alexander Shapiro. 2003. On a class of nonsmooth composite functions. *Mathematics of Operations Research* 28, 4 (2003), 677–692. https://doi.org/10.1287/moor.28.4.677.20512
[10] Christopher J. C. H. Watkins and Peter Dayan. 1992. Q-Learning. *Machine Learning* 8, 3 (01 May 1992), 279–292. https://doi.org/10.1007/BF00992698
[11] Lu Yu, Krishnakumar Balasubramanian, Stanislav Volgushev, and Murat A Erdogdu. 2021. An Analysis of Constant Step Size SGD in the Non-convex Regime: Asymptotic Normality and Bias. In *Advances in Neural Information Processing Systems*, M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan (Eds.), Vol. 34. Curran Associates, Inc., 4234–4248. https://proceedings.neurips.cc/paper_files/paper/2021/file/21ce689121e39821d07d04faab328370-Paper.pdf
[12] Yixuan Zhang, Dongyan Huo, Yudong Chen, and Qiaomin Xie. 2024. Prelimit Coupling and Steady-State Convergence of Constant-stepsize Nonsmooth Contractive SA. arXiv:2404.06023 [stat.ML]