Talk Less, Work More: Communication-Efficient Decentralized Stochastic Approximation

Tianyu Cao **Haixiang Sun** **Yang Xu**
Edwardson School of Industrial Engineering
Purdue University
{cao357, sun1321, xu1720}@purdue.edu

Abstract

Decentralized stochastic approximation (SA) enables cooperative fixed-point solving over networks without a central server. Local updates are often inexpensive relative to communication, making frequent information exchange a bottleneck. We propose a decentralized SA method in which each agent performs H local updates before each communication round. This method reduces communication frequency and can be combined with FastMix to improve consensus on poorly connected graphs. We establish finite-time bounds under contraction of the averaged operator in a general norm and Markovian sampling, using a Lyapunov function that tracks estimation error, consensus error, and delayed terms. The bounds characterize tradeoffs among local computation, communication, and sampling. Numerical experiments on decentralized Q-learning, linear SA, and Markovian SGD illustrate these tradeoffs.

1 Introduction

Stochastic Approximation (SA) is a fundamental approach for solving fixed point and root finding problems when the underlying operator is only observed through noisy samples [25, 20]. It has applications in stochastic optimization [6], reinforcement learning [32], game theory [16], and dynamical systems [19].

In modern deployments, data are distributed across multiple agents connected by a network without a central coordinator. This motivates decentralized stochastic approximation [4], where m agents cooperatively solve a fixed point equation over a mesh network [31], whose target is the fixed point of the averaged mean operator

$$x = \bar{F}(x) := \frac{1}{m}\sum_{i=1}^{m} \bar{F}_i(x), \tag{1}$$

where $\bar{F}_i \colon \mathbb{R}^d \to \mathbb{R}^d$ is the local mean operator accessed by agent i through noisy samples. Existing approaches alternate one stochastic approximation update with one communication step [18, 41]. In practice, communication cost is expensive compared with local computation. Full-precision update exchange in large-scale distributed training could be sizable per iteration, making communication a bottleneck [1, 22, 5]. The bottlenecks of communication also arise in federated learning, where limited device bandwidth and energy budgets, together with intermittent connectivity, further hinder scalability [17]. To reduce communication overhead, local-update strategies have been widely adopted, allowing agents to perform multiple local steps between communication rounds. Representative examples of local-update methods include local SGD [30] and its federated counterparts [23]. By performing multiple local steps between communication rounds, these methods can substantially reduce synchronization frequency compared with fully synchronous training.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

Table 1: Comparison of settings for decentralized stochastic approximation. Multiple local steps mean $H > 1$. Average-only contraction allows noncontractive local mean operators. Non-asymptotic analysis means finite-time analysis. Check marks, crosses, and dashes mean covered, not covered, and not applicable. FastMix is not applicable to centralized LSA.

Work	Decentralized	Multi-local-step	Markovian	General norm	Avg-only contraction	Non-asymptotic	Fast mix
DSA [4]	✓	✗	✗	✗	✓	✗	✗
Biased DSA [33]	✓	✗	✓	✗	✓	✓	✗
DCSA [41]	✓	✗	✓	✗	✓	✓	✗
LSA [14]	✗	✓	✓	✗	✓	✓	–
DLMSA [34]	✓	✓	✓	✓	✗	✓	✗
DLSA (ours)	✓	✓	✓	✓	✓	✓	✓

Terminology. Here "average-only contraction" means that only the averaged operator $\bar{F}$ is assumed contractive. The local mean operators $\bar{F}_i$ may be noncontractive. We use "finite-time" and "non-asymptotic" synonymously.

In decentralized training with no central node for aggregation, multiple communication steps are usually required in each communication round, especially when the network connectivity is poor. In all existing works on decentralized stochastic approximation [4, 33, 41, 34], only a single gossip step is adopted per communication round, making their decentralized algorithms not robust to poorly-connected networks. Beyond communication bottlenecks, another challenge in solving (1) is that the averaged operator $\bar{F}$ is contractive only under a problem-dependent, non-Euclidean norm. Examples arise in tabular reinforcement learning, where Bellman operators are naturally contractive in the ℓ_∞ norm [11]. Related analyses using problem-dependent norms also appear in robust average-reward reinforcement learning [37, 36]. However, existing local stochastic approximation results do not yet provide a complete theory for fully decentralized settings under non-Euclidean norm contractions as shown in Table 1.

In this work, we develop a communication-efficient decentralized stochastic approximation method with local updates that explicitly addresses these issues. Our Decentralized Local Stochastic Approximation (DLSA) algorithm targets the setting in which the averaged operator is contractive with respect to an arbitrary norm, while allowing multiple local updates. DLSA organizes communication at two levels. The higher level is communication round which happens after every H local updates. A large H can save communication costs by reducing the communication frequency. Within each communication round, an optional accelerated multi-consensus routine based on FastMix improves mixing when a single gossip step is insufficient on a poorly connected network. This perspective allows us to study communication efficiency rather than only iteration complexity, which is the dominant focus of most existing works.

On the theory side, we use a generalized Moreau envelope to construct a potential function compatible with the norm in the contraction assumption and embed it into a Lyapunov function that also tracks consensus error and a delayed term induced by Markovian sampling. Our finite-time analysis jointly handles the errors introduced by multiple local updates, contraction of only the averaged operator under a general norm, Markovian sampling, and consensus. The results demonstrate tradeoffs among communication, computation and sampling: i) large H reduces communication frequency and cost, but may enlarge the ultimate error ball under steady state, ii) better network connectivity permits slower mixing of the Markovian samples.

Contributions

- We propose DLSA, a decentralized stochastic approximation algorithm that performs multiple local SA updates per communication round, yielding a direct reduction in communication frequency while approaching the target fixed point.
- For poorly-connected network where one-step gossip is insufficient in each communication round, we adopt an accelerated multi-consensus routine based on FastMix to enhance the communication, making our algorithm robust to the quality of network.
- We develop a finite time analysis for DLSA when only the averaged operator is contractive under an arbitrary norm, covering non Euclidean contractions such as the infinity norm in tabular Q-learning or some nontrivial norm in averaged reward TD learning.

- We evaluate DLSA on various settings, including linear stochastic approximation, Markovian stochastic gradient descent and reinforcement learning, and demonstrate substantial communication savings.

2 Related works

Distributed stochastic approximation typically combines local stochastic updates with gossip communications [8], so that all agents approach a common fixed point of the averaged operator with converged consensus error. An asymptotic convergence analysis of decentralized SA with i.i.d. distribution of the noise in operator is derived in [4], while finite-time theory of decentralized SA under Markovian noise is more preferred for practice. Non-asymptotic analysis under Markovian noise was relatively underexplored even in the centralized setting, until recently: [3] and [29] study linear SA, while [24, 12] focus on nonlinear SA.

Building on these developments, finite-time analysis of decentralized SA with Markovian noise has been studied in [13, 15, 21, 41, 34]. Specifically, [13, 15, 21] study linear decentralized SA in the context of temporal-difference learning. The closest nonlinear references are [41] and [34]. The former studies one-local-step-one-communication decentralized SA in Euclidean geometry. The latter studies the same local-update template as ours, but under the substantially stronger assumption that every local mean operator $\bar{F}_i$ is contractive. DCSA [41] uses one SA update and one gossip step per round. DLSA with $H = 1$ has the same per-round resource counts but a different update order. DCSA adds a local SA increment to the mixed previous iterates, whereas DLSA mixes the locally updated iterates. Comparisons across H within DLSA concern local computation, while plain-gossip/FastMix comparisons concern consensus acceleration. Table 1 compares theoretical settings.

Most existing decentralized SA results focus on iteration complexity rather than communication complexity [9, 27]. Some works discuss communication costs, especially in federated settings [40, 28], but in decentralized computing with poorly connected network, multi-consensus mixing in each communication round is usually required to improve the network connectivity. The communication complexity additionally depends on the connectivity of networks and is hence not equivalent to iteration complexity. Polynomial-gossip schemes such as FastMix [38, 39] and Chebyshev acceleration [2, 26] can accelerate the multi-consensus routine and improve the network dependence. To the best of our knowledge, the decentralized SA works in Table 1 use plain gossip or centralized averaging [4, 33, 41, 14, 34], non-robust to poorly connected networks. Our contribution is to apply FastMix to decentralized local SA with Markovian sampling and general contractive norms, as well as provide the analysis on communication complexity.

3 Problem settings

This work studies finite-sample decentralized stochastic approximation over a network of m agents connected by a static, connected, undirected graph, without a central coordinator. The goal is to solve for the fixed point $x^\star \in \mathbb{R}^d$ of the averaged operator

$$x^\star = \bar{F}(x^\star), \qquad \bar{F}(x) := \frac{1}{m}\sum_{i=1}^{m} \bar{F}_i(x), \tag{2}$$

where $\bar{F}_i(x) := \mathbb{E}_{Y_i \sim \mu_i}[F_i(x, Y_i)]$, $F_i : \mathbb{R}^d \times \mathcal{Y} \to \mathbb{R}^d$, and $\mathcal{Y}$ denotes sample space with distribution μ_i. For each agent $i \in \{1, \cdots, m\}$, let $\{Y_i^{(s)}\}_{s\geq 0}$ denote the underlying Markov chain. The sampling chains have fixed transition kernels that do not depend on the iterates. Conditioning on the joint past does not change these kernels. During round k and inner step t, Algorithm 1 uses the sample $Y_{i,t}^k := Y_i^{(kH+t)}$. For any $\tilde{\epsilon} > 0$, denote by $\tilde{\tau}_{\tilde{\epsilon}}$ the smallest integer such that for all $y_i \in \mathcal{Y}$,

$$d_{TV}\left(P\left(Y_i^{(s)} = \cdot \mid Y_i^{(0)} = y_i\right), \mu_i\right) \leq \tilde{\epsilon}, \qquad \forall s \geq \tilde{\tau}_{\tilde{\epsilon}}. \tag{3}$$

Here d_{TV} denotes total variation distance. In addition, we assume $\mathcal{Y}$ is finite and each chain $\{Y_i^{(s)}\}_{s\geq 0}$ is irreducible and aperiodic. Then there exist constants $r_p \in (0, 1)$ and $c_p > 0$ such that

$$d_{TV}\left(P\left(Y_i^{(s)} = \cdot \mid Y_i^{(0)} = y_i\right), \mu_i\right) \leq c_p r_p^s, \qquad \forall s \geq 0, \tag{4}$$

and therefore $\tilde{\tau}_{\tilde{\epsilon}} \leq \max\{0, \lceil \tau_{\tilde{\epsilon}} \rceil\}$, where $\tau_{\tilde{\epsilon}} := \frac{1}{1-r_p} \log \frac{c_p}{\tilde{\epsilon}}$.

Besides, we have the following assumptions on the operator and network.

Assumption 1 (Operator)**.** *Let $\|\cdot\|_c$ denote a general norm satisfying the following conditions.*

(i) There exists $L > 0$, such that for any $x_1, x_2 \in \mathbb{R}^d, y \in \mathcal{Y}$, we have

$$\max\{\|F_i(x_1, y) - F_i(x_2, y)\|_c, \|F_i(x_1, y) - x_1 - F_i(x_2, y) + x_2\|_c\} \leq L\|x_1 - x_2\|_c, \quad (5)$$

(ii) There exists some $\beta \in (0, 1)$ such that for any $x, y \in \mathbb{R}^d$,

$$\|\bar{F}(x) - \bar{F}(y)\|_c \leq \beta\|x - y\|_c. \quad (6)$$

(iii) There exists a universal constant $B > 0$, such that for any $i \in [m], y \in \mathcal{Y}$,

$$\|F_i(x^\star, y)\|_c + \|x^\star\|_c \leq B. \quad (7)$$

Assumption 1 (ii) guarantees that $x^\star$ exists and is unique.

Remark 1 (Why average-only contraction is weaker)**.** *Consider two agents with $x = (x_1, x_2)^\top \in \mathbb{R}^2$ and local quadratic losses*

$$f_1(x) = \tfrac{1}{2}(x_1 - 1)^2, \qquad f_2(x) = \tfrac{1}{2}(x_2 + 1)^2.$$

Their average has the unique minimizer $x^\star = (1, -1)^\top$. The mean operators $\bar{F}_i(x) = x - \nabla f_i(x)$ are

$$\bar{F}_1(x) = (1, x_2)^\top, \qquad \bar{F}_2(x) = (x_1, -1)^\top.$$

Neither local operator is a strict Euclidean contraction, since it leaves one coordinate unchanged. Their average satisfies

$$\bar{F}(x) = \tfrac{1}{2}x + \tfrac{1}{2}(1, -1)^\top, \qquad \|\bar{F}(x) - \bar{F}(y)\|_2 = \tfrac{1}{2}\|x - y\|_2.$$

Thus contraction of the averaged operator holds even when neither local operator is strictly contractive.

Assumption 2 (Network)**.** *For the static connected undirected graph, the nonnegative matrix $W = (w_{ij}) \in \mathbb{R}^{m \times m}$ satisfies the following conditions.*

(i) Compliance with graph: For $i \neq j$, the weights are positive on graph edges and zero otherwise. The self-weights satisfy $w_{ii} > 0$.

(ii) Symmetry: $W^\top = W$.

(iii) Doubly stochastic: $W\mathbf{1} = \mathbf{1}$, $\mathbf{1}^\top W = \mathbf{1}^\top$, where $\mathbf{1} \in \mathbb{R}^m$ is an all-one vector.

The weight matrix W is usually constructed using the Metropolis rule [35]. Denote the network parameter by $\rho = \|W - \frac{1}{m}\mathbf{1}\mathbf{1}^\top\|_2 \in [0, 1)$. A larger ρ means slower mixing over the network. To obtain a smaller ρ for a given graph, an offline semidefinite program can minimize ρ subject to the graph constraints [7].

Notations: Use $[m]$ to denote $\{1, \cdots, m\}$. Stacked variables are column vectors $X = [x_1^\top, \ldots, x_m^\top]^\top \in \mathbb{R}^{md}$, where $x_i \in \mathbb{R}^d$ is the local iterate of agent $i \in [m]$. Denote the averaged variable $\bar{x} := \frac{1}{m}\sum_{i=1}^m x_i$. Denote I_d as the identity matrix with dimension d. Use $\otimes$ to denote the Kronecker product. The Euclidean norm controls consensus error, while $\|\cdot\|_c$ is used for the contraction assumption and the final error bounds. Use u_{2c} and ℓ_{2c} to denote norm equivalence between $\|.\|_2$ and $\|.\|_c$: for any $x \in \mathbb{R}^d$, it holds $\ell_{2c}\|x\|_2 \leq \|x\|_c \leq u_{2c}\|x\|_2$ and $\delta := \frac{u_{2c}}{\ell_{2c}}$.

4 Algorithm

DLSA alternates local stochastic approximation and communication, as shown in Algorithm 1. At the beginning of round k, agent i starts from x_i^k and performs H updates using its local sampled operator F_i. The agents then communicate the resulting iterates with their neighbors.

Algorithm 1 Decentralized Local Stochastic Approximation (DLSA)

Input $X^0 = (x_i^0)_{i\in[m]}$
Parameters stepsize $\alpha > 0$, local-update length $H \in \mathbb{N}$, number of rounds K, communication operator $\mathcal{C}$
for $k = 0, \dots, K-1$ **do**
  $z_{i,0}^k = x_i^k$ for every $i \in [m]$
  for $t = 0, \dots, H-1$ **do**
    $z_{i,t+1}^k = z_{i,t}^k + \alpha\big(F_i(z_{i,t}^k, Y_{i,t}^k) - z_{i,t}^k\big)$ for every $i \in [m]$
  end for
  $X^{k+1} = \mathcal{C}(Z_H^k)$
end for
Return X^K

Algorithm 2 FastMix$(Z^0, W_b, n_\rho, \eta_w)$

Input stacked iterate $Z^0 \in \mathbb{R}^{md}$, base $0 \preceq W_b \preceq I$, number of gossip steps n_ρ, momentum parameter η_w
Initialize $Z^1 = Z^0$
for $\nu = 1, \dots, n_\rho$ **do**
  $Z^{\nu+1} = (1+\eta_w)(W_b \otimes I_d)Z^\nu - \eta_w Z^{\nu-1}$
end for
Return $Z^{n_\rho+1}$

Write $J = \mathbf{1}\mathbf{1}^\top/m$. For plain DLSA, the communication operator is $\mathcal{C}(Z_H^k) = (W \otimes I_d)Z_H^k$, which requires one gossip step. When the network is poorly connected, we instead use DLSA-FastMix with a positive-semidefinite base matrix W_b. We take $W_b = W$ if $W \succeq 0$, and $W_b = (I+W)/2$ otherwise, so that $0 \preceq W_b \preceq I$ and $\rho_b := \|W_b - J\|_2 < 1$. The latter choice preserves the graph and averaging property, and each application requires one neighbor exchange with W followed by local averaging. We set

$$\mathcal{C}(Z_H^k) = \text{FastMix}(Z_H^k, W_b, n_\rho, \eta_w),$$
$$\eta_w = \frac{1}{1+\sqrt{1-\rho_b^2}}.$$

This base satisfies the assumptions of Proposition 1 in [38]. Plain DLSA continues to use the original matrix under Assumption 2.

Repeated multiplication by W contracts disagreement by ρ^{n_ρ} and needs $\mathcal{O}((1-\rho)^{-1})$ steps to reach a fixed threshold. FastMix instead produces the average-preserving polynomial filter $\widetilde{W} = P_{n_\rho}(W_b)$. Its output is $Z^{n_\rho+1} = (\widetilde{W} \otimes I_d)Z^0$. Let $c_{\text{fm}} := 1 - 1/\sqrt{2}$, $\bar{z}^0 = m^{-1}\sum_i z_i^0$, and $D_\nu := \|Z^\nu - \mathbf{1} \otimes \bar{z}^0\|_2$. With $\rho_{\text{fm}} := \|\widetilde{W} - J\|_2$, the cited proposition yields

$$D_{n_\rho+1} \le \rho_{\text{fm}} D_0, \tag{8}$$
$$\rho_{\text{fm}} \le \sqrt{14}\big(1 - c_{\text{fm}}\sqrt{1-\rho_b}\big)^{n_\rho}. \tag{9}$$

Remark 2. *For a target $\bar{\rho} \in (0,1)$, it suffices to choose the integer n_ρ with*

$$n_\rho \ge \frac{\log(\sqrt{14}/\bar{\rho})}{c_{\text{fm}}\sqrt{1-\rho_b}}.$$

Then $\rho_{\text{fm}} \le \bar{\rho}$. If $W_b = W$, then $\rho_b = \rho$. Otherwise $\rho_b \le (1+\rho)/2$. Hence $1-\rho_b \ge (1-\rho)/2$ in either case. Each base multiplication costs one gossip exchange, so the sufficient communication cost remains $\mathcal{O}((1-\rho)^{-1/2})$ up to logarithmic factors.

4.1 Main Results

For convenience, define the consensus error $S_k := \frac{1}{m}\sum_{i=1}^m \|x_i^k - \bar{x}^k\|_c$ and $S_{k,t} := \frac{1}{m}\sum_{i=1}^m \|z_{i,t}^k - \bar{z}_t^k\|_c$, and the estimation gap $R_k := \big\|\bar{x}^k - x^\star\big\|_c$ and $R_{k,t} := \|\bar{z}_t^k - x^\star\|_c$. To control the bias from Markovian sampling, we use a delay of τ communication rounds. Each round contains H Markov

transitions, so τ is chosen such that $H\tau$ is at least the mixing time $\tilde{\tau}_\epsilon$ measured in transitions, with the precise choice given in Appendix A.3. We now show the convergence results of Algorithm 1 in Theorem 1, with parameter details given in Appendix A.4.

Theorem 1 (Finite-time convergence of DLSA)**.** *Suppose Assumptions 1 and 2 hold, and suppose the Markovian sampling satisfies the mixing condition* (4)*. Run Algorithm 1 with constant stepsize α and local-update length H. Let ρ_{eff} denote the effective network mixing factor in one communication round. For plain DLSA, $\rho_{\text{eff}} = \rho$; for DLSA-FastMix, $\rho_{\text{eff}} = \rho_{\text{fm}}$. Suppose*

$$\alpha \leq \frac{c_\alpha}{H(\tau+1)(1+L)^2} = \widetilde{O}\left(\frac{1}{H\tau(1+L)^2}\right), \tag{10}$$

$$\rho_{\text{eff}} \leq \frac{c_\rho}{\sqrt{m}\,(\tau+1)} = \widetilde{O}\left(\frac{1}{\sqrt{m}\tau}\right), \tag{11}$$

where $c_\alpha, c_\rho > 0$ are problem-dependent constants, $\widetilde{\mathcal{O}}(\cdot)$ hides logarithmic factors and some problem-dependent constants, τ denotes Markov delay and satisfies $\tau = \widetilde{\mathcal{O}}\Big(1 + \frac{1}{H(1-r_p)}\Big)$, then for any $k \geq \tau$,

$$\begin{aligned}\mathbb{E}\big[R_k^2\big] \leq & C_R(1-c_0\alpha H)^{k-\tau}\big(\mathbb{E}\big[R_\tau^2\big]+\mathbb{E}\big[S_\tau^2\big]+\mathbb{E}\big[S_0^2\big]\big) \\ & +\widetilde{\mathcal{O}}\big(\alpha B^2\big(1+H+\alpha H^2(1+L)^2+\alpha^2H^3(1+L)^2\big)\big),\end{aligned} \tag{12}$$

$$\begin{aligned}\mathbb{E}\big[S_k^2\big] \leq & \, C_S\left(1-c_0\alpha H\right)^{k-\tau}\left(\mathbb{E}\big[R_\tau^2\big]+\mathbb{E}\big[S_\tau^2\big]+\mathbb{E}\big[S_0^2\big]\right) \\ & +\widetilde{\mathcal{O}}\left(\rho_{\text{eff}}^2\alpha^2H^2B^2\left(1+\alpha(1+L)^2\left(1+H+\alpha H^2(1+L)^2+\alpha^2H^3(1+L)^2\right)\right)\right),\end{aligned} \tag{13}$$

where $c_0, C_R, C_S > 0$ are constants independent of α, H, B and ρ_{eff}.

The error bound on the right-hand side uses initial errors at the delayed starting round because the Lyapunov recursion requires a full history of past iterates before it can begin. Theorem 1 shows that both the estimation error R_k^2 and the consensus error S_k^2 contract geometrically up to steady-state error terms. The result highlights two key tradeoffs.

(i) The theorem makes the tradeoff in H explicit: larger H improves the transient contraction factor through $\alpha\kappa H$, but the error ball typically grows polynomially with H. Thus larger H can enlarge the steady-state error ball unless the stepsize is reduced accordingly.

(ii) Algorithm 1 actually involves two mixing processes: one is consensus mixing through the physical network with quality measured by ρ_{eff}; another one is the mixing of the Markovian sampling process, whose quality is reflected by τ (small τ is wanted). The condition (11) implicitly implies a trade-off relationship between these two mixings: the better quality network tolerates a worse sampling process, and vice versa.

We use communication complexity to count gossip steps and sample complexity to count samples per agent. Thus, K communication rounds require K gossip steps for plain DLSA or Kn_ρ for DLSA-FastMix, and each agent uses KH samples. To achieve ε-precision for estimation error and consensus error, we have following results on communication and sample complexities, by simply applying Theorem 1.

Corollary 1 (Communication and sample complexity)**.** *Fix some small enough $\varepsilon \in (0,1)$. With the same setting as in Theorem 1, denote $\alpha_{\max} := \frac{c_\alpha}{H(\tau+1)(1+L)^2}$. Choose an admissible constant stepsize on the scale $\min\left\{\alpha_{\max}, \frac{\varepsilon}{B^2(1+H)}, \frac{\sqrt{\varepsilon}}{BH}\right\}$ with sufficiently small multiplicative constants and logarithmic corrections so that both steady-state terms in Theorem 1 are at most $\varepsilon/2$. Then, after $K = \widetilde{\mathcal{O}}\left(\tau + \frac{1}{\alpha H}\log\frac{1}{\varepsilon}\right)$ communication rounds, we obtain*

$$\mathbb{E}\big[R_K^2\big] \leq \varepsilon, \quad \mathbb{E}\big[S_K^2\big] \leq \varepsilon$$

For plain DLSA, the communication and sample complexities are respectively

$$\widetilde{\mathcal{O}}\left(\tau + \max\left\{(\tau+1)(1+L)^2, \frac{B^2(1+H)}{H\varepsilon}, \frac{B}{\sqrt{\varepsilon}}\right\}\log\frac{1}{\varepsilon}\right), \tag{14}$$

and

$$\widetilde{\mathcal{O}}\left(\tau H + \max\left\{H(\tau+1)(1+L)^2, \frac{B^2(1+H)}{\varepsilon}, \frac{BH}{\sqrt{\varepsilon}}\right\}\log\frac{1}{\varepsilon}\right). \tag{15}$$

FastMix is used when the network is poor (ρ does not meet (11)*), the communication cost is*

$$\widetilde{\mathcal{O}}\left(\frac{1}{\sqrt{1-\rho}}\left(\tau + \max\left\{(\tau+1)(1+L)^2, \frac{B^2(1+H)}{H\varepsilon}, \frac{B}{\sqrt{\varepsilon}}\right\}\right)\log\frac{1}{\varepsilon}\right).$$

The effectness of H. More local updates strengthen contraction within each communication round at a fixed stepsize. They can also enlarge the error floor, and the admissible stepsize may need to decrease. The bound therefore describes a tradeoff rather than a monotone improvement with more local computation. Moderate local-update lengths can reduce communication over the regimes illustrated by the experiments.

Q-learning is a central application of stochastic approximation in reinforcement learning. Since the Bellman optimality operator is naturally contractive under the ℓ_∞ norm, Q-learning serves as a representative example of the general-norm setting addressed by Theorem 1.

4.2 Application in Decentralized Q-learning

We specialize the general decentralized SA framework to fixed-behavior tabular decentralized Q-learning. Let $\mathcal{S}$ and $\mathcal{A}$ be finite, fix $\gamma_{\text{disc}} \in (0,1)$, and for each agent $i \in [m]$ let $r_i : \mathcal{S}\times\mathcal{A} \to [-R_{\max}, R_{\max}]$ be a deterministic reward function and $P_i(\cdot \mid s,a)$ its local transition kernel. Each agent follows a fixed behavior policy $\pi_i(\cdot \mid s)$, independent of the current Q-iterate. Transitions do not depend on other agents' states or actions. The independent local sample streams have irreducible and aperiodic state-action chains, and agents communicate only their Q-tables. The local sample stream is $Y_i^{(\ell)} := \big(S_i^{(\ell)}, A_i^{(\ell)}, R_i^{(\ell)}, S_i^{(\ell+1)}\big)$, with $R_i^{(\ell)} = r_i\big(S_i^{(\ell)}, A_i^{(\ell)}\big)$ and $S_i^{(\ell+1)} \sim P_i(\cdot \mid S_i^{(\ell)}, A_i^{(\ell)})$, and stationary distribution $\mu_i^{\pi_i}$. The corresponding Bellman optimality operator is in the form of

$$(T_iQ)(s,a) := r_i(s,a) + \gamma_{\text{disc}} \sum_{s'\in\mathcal{S}} P_i(s' \mid s,a) \max_{a'\in\mathcal{A}} Q(s',a'). \tag{16}$$

Let $d := |\mathcal{S}||\mathcal{A}|$ and let $e_{s,a} \in \mathbb{R}^d$ be the coordinate basis vector associated with (s,a). For a sample $y = (s,a,r,s')$, define the sampled Bellman update operator

$$F_i(Q,y) := Q + e_{s,a}\Big(r + \gamma_{\text{disc}} \max_{a'\in\mathcal{A}} Q(s',a') - Q(s,a)\Big). \tag{17}$$

With the local samples $Y_{i,t}^k := Y_i^{(kH+t)}$, the inner loop of Algorithm 1 reduces to standard tabular Q-learning updates performed locally at each agent. Let $r_{p,Q}$ denote the Markovian mixing rate induced by the behavior policy and the environment.

We now identify the correct mean operator seen by the decentralized stochastic approximation setting. Let $d_i(s,a) := P_{\mu_i^{\pi_i}}\big((S_i,A_i) = (s,a)\big)$, $\bar{d}(s,a) := \frac{1}{m}\sum_{i=1}^m d_i(s,a)$, and $d_{\min} := \min_{(s,a)\in\mathcal{S}\times\mathcal{A}} \bar{d}(s,a)$, and assume $d_{\min} > 0$. Define $\bar{D} := \operatorname{diag}(\bar{d}(s,a) : (s,a) \in \mathcal{S}\times\mathcal{A})$, together with the coverage-weighted Bellman operator

$$(\widetilde{T}Q)(s,a) := \frac{\sum_{i=1}^m d_i(s,a)\,(T_iQ)(s,a)}{\sum_{i=1}^m d_i(s,a)}. \tag{18}$$

Appendix B shows that the averaged mean operator of the decentralized SA model has the form of $\bar{F}_Q(Q) = Q + \bar{D}(\widetilde{T}Q - Q)$ with Assumption 1 guaranteed with $\|\cdot\|_c = \|\cdot\|_\infty$ and constants $L_Q = 1+\gamma_{\text{disc}}$, $\beta_Q = 1-(1-\gamma_{\text{disc}})d_{\min}$ and $B_Q = \frac{2R_{\max}}{1-\gamma_{\text{disc}}}$. Thus, the fixed point coincides with the unique fixed point $Q^\star$ of $\widetilde{T}$. $Q^\star$ is the optimal Q-function of an aggregate MDP with weights $d_i(s,a)/\sum_j d_j(s,a)$, reducing to uniform averaging under identical coverage. Its greedy policy, which selects an action maximizing $Q^\star(s,a)$ at each state, need not be optimal for an individual agent. The Q-function induced by this policy in a local MDP can differ from $Q^\star$, as quantified in Appendix B.5. We can derive the following corollary:

Corollary 2 (Decentralized Q-learning specialization)**.** *Assume the tabular Q-learning setting above, the network assumption, and geometric Markov mixing* (4) *hold. Run Algorithm 1 with Bellman update F_i defined above, constant stepsize α, and local length H.*

Let τ_Q and $\kappa_Q > 0$ be the Q-learning delay and Moreau-envelope contraction quantity defined in Appendix B. Suppose the stepsize and effective network factor satisfy the admissibility conditions of Theorem 1 after substituting the Q-learning constants, and define $\mathcal{P}_Q(\alpha, H) := 1 + H + \alpha H^2(1 + L_Q)^2 + \alpha^2 H^3(1 + L_Q)^2$. Then, for all $k \geq \tau_Q$,

$$\mathbb{E}\big[R_k^2\big] \leq \widetilde{\mathcal{O}}\left(\left(1 - \frac{\alpha\kappa_Q H}{32}\right)^{k-\tau_Q}\right) + \widetilde{\mathcal{O}}\left(\alpha B_Q^2\, \mathcal{P}_Q(\alpha, H)\right), \tag{19}$$

$$\mathbb{E}\big[S_k^2\big] \leq \widetilde{\mathcal{O}}\left(\left(1 - \frac{\alpha\kappa_Q H}{32}\right)^{k-\tau_Q}\right) + \widetilde{\mathcal{O}}\left(\rho_{\text{eff}}^2 \alpha^2 H^2 B_Q^2 \left(1 + \alpha(1 + L_Q)^2 \mathcal{P}_Q(\alpha, H)\right)\right). \tag{20}$$

For fixed H and problem parameters, an admissible stepsize $\alpha = \Theta(\varepsilon)$ gives the numbers of communication rounds and samples per agent

$$\widetilde{\mathcal{O}}\left(\tau_Q + \frac{1}{\kappa_Q H \varepsilon} \log \frac{1}{\varepsilon}\right) \quad \textit{and} \quad \widetilde{\mathcal{O}}\left(\tau_Q H + \frac{1}{\kappa_Q \varepsilon} \log \frac{1}{\varepsilon}\right),$$

respectively. Plain DLSA uses one gossip step per round, while DLSA-FastMix uses n_ρ gossip steps per round. Thus its communication complexity is n_ρ times the round count.

4.3 Analysis sketch

The proof tracks the estimation error of the averaged iterate and the disagreement among agents. A generalized Moreau envelope provides a smooth potential compatible with the general norm in the contraction assumption [10]. This potential allows us to use the contraction of the averaged mean operator even when the individual operators are not contractive.

Multiple local updates can increase disagreement because agents use different operators and samples. The inner-loop recursions bound this increase using the errors at the beginning of each round. The communication phase then contracts disagreement according to the effective network mixing factor. These two bounds quantify the competing effects of longer local computation and stronger consensus.

Markovian samples require an additional comparison with the past. Conditioning on the history from τ rounds earlier leaves at least $H\tau$ sample transitions for mixing, so the conditional sampling bias is small. Comparing the current iterate with a delayed iterate separates this bias from the changes caused by local updates. The resulting estimates are derived in Appendix A.3.

Finally, we combine the smooth potential with current and delayed consensus errors in a Lyapunov function. Under the stepsize and network conditions, the positive error terms can be absorbed into its decrease. This gives geometric convergence up to a neighborhood whose size depends on the stepsize, local-update length, and network mixing factor. Appendix A.4 gives the full recursion and the error bounds used in Theorem 1.

5 Experiments

In this section, we evaluate DLSA on decentralized Q-learning. These synthetic experiments provide controlled numerical validation of the communication tradeoff on a poorly connected path graph. Following Algorithm 1, one outer round consists of H local Q-learning updates followed by one communication phase. The local-update comparisons concern communication frequency, while the plain-gossip/FastMix comparisons concern consensus acceleration. We use a constant stepsize within each run, with α_H selected separately across H to obtain comparable terminal errors. This design does not isolate the error-ball dependence on H under a common stepsize. For clearer visualization, we plot ℓ_2 error in all figures.

We first consider a decentralized Q-learning problem with discount factor $\gamma = 0.95$, whose Bellman optimality operator is naturally contractive under the non-Euclidean ℓ_∞ norm. We generate $m = 10$ agents, each with its own MDP, and connect them through a fixed path graph with a doubly-stochastic

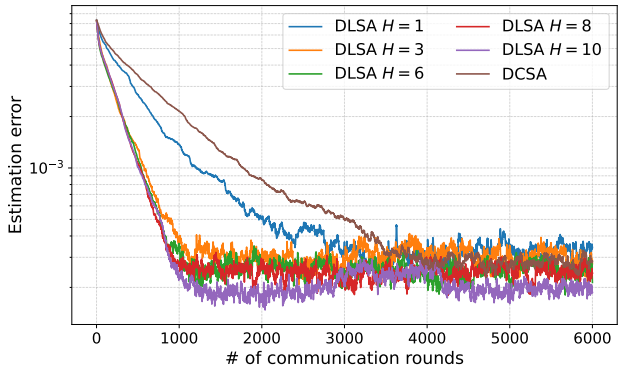


Figure 1: Estimation error analysis

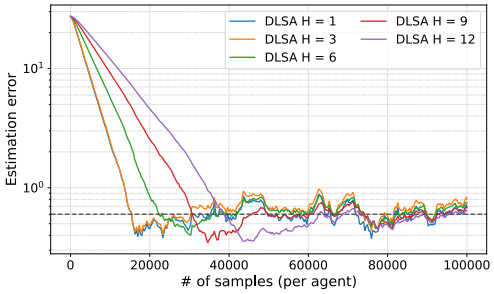


Figure 2: Sample complexity analysis

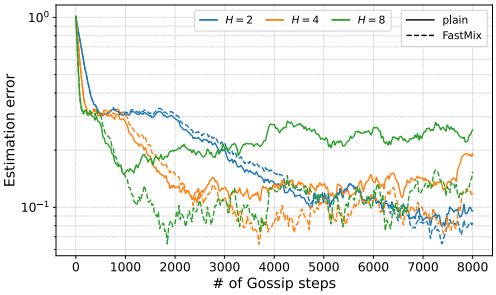


Figure 3: FastMix comparison

mixing matrix W given by Metropolis weights, i.e., $w_{ij} = 1/(1 + \max\{\deg(i), \deg(j)\})$ for each edge (i, j) and $w_{ii} = 1 - \sum_{j \neq i} w_{ij}$. Here $\deg(i)$ is the degree of node i. Each agent samples from its own MDP and maintains a local Q-table $Q_i \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{A}|}$. We compute a reference solution $Q^\star$ as the fixed point of the averaged Bellman optimality operator and report the network-average table error

$$\mathrm{err}_k := \frac{1}{m} \sum_{i=1}^{m} \|Q_{i,k} - Q^\star\|_2.$$

Figure 1 reports the estimation error against communication rounds for DLSA with $H \in \{1, 3, 6, 8, 10\}$ and a separately labelled DCSA baseline. Each DLSA round contains H local updates followed by one gossip step. Within the DLSA curves, the larger local-update lengths reach a common error level earlier than the $H = 1$ run. DCSA [41] is a one-update-per-round baseline. Its local SA increment is added to the mixed previous iterates, while DLSA with $H = 1$ mixes the updated iterates. Thus the two methods have the same per-round resource counts but different recursions.

Figure 2 gives a separate comparison against samples per agent for $H \in \{1, 3, 6, 9, 12\}$. In this plot, moderate local-update lengths such as $H = 3$ and $H = 6$ remain competitive in samples, while $H = 9$ and $H = 12$ take more samples to reach the displayed error levels. These curves illustrate that savings in communication need not yield the same savings in samples. We interpret the comparisons within each figure and do not compare absolute error levels across Figures 1 and 2.

We also test DLSA-FastMix on a poorly-connected path graph with $m = 8$ agents over heterogeneous random MDPs, where $\gamma = 0.9$ and $H \in \{2, 4, 8\}$. We compare DLSA-FastMix with $n_\rho = 12$ polynomial-filter gossip steps per round against plain DLSA with one gossip step per round, and plot the per-agent estimation error against the total number of gossip steps. The step sizes are tuned separately for different values of H, and for each fixed H, DLSA and DLSA-FastMix use the same stepsize. As shown in Figure 3, FastMix reaches a lower steady-state error than plain gossip for every H, confirming that the accelerated multi-consensus routine improves the effective mixing on

poorly-connected graphs. Moreover, the gap between FastMix and plain gossip becomes larger as H increases, which is consistent with Theorem 1: larger H amplifies the effect of the consensus error, and thus a smaller effective network mixing factor ρ_{eff} becomes more important.

Additional experiments on linear stochastic approximation and stochastic gradient descent with Markovian noise are reported in Appendix C. The same appendix gives the MDP generation details and the stepsize-sensitivity study.

6 Conclusion

In this work, we proposed DLSA, a communication-efficient decentralized stochastic approximation method that performs multiple local SA updates between successive communication steps, thereby substantially reducing the number of communication rounds required to approach the desired fixed point, and further enabling additional gains when combined with communication acceleration techniques. On the theoretical side, we established a finite-time convergence guarantee for DLSA under the assumption that the averaged operator is contractive under an arbitrary norm, which allows our analysis to handle non-Euclidean contractions, including $\|\cdot\|_\infty$-type settings arising in tabular Q-learning. The experiments on Q-learning, linear SA, and Markovian SGD provide controlled illustrations of the communication tradeoff.

Despite its advantages, our theory suggests that using a larger number of local updates H may amplify the resulting error ball, and hence requires additional stepsize tuning to ensure the final accuracy meets a target precision. Our guarantees require fixed H, α, stationary sampling, and static, synchronous, reliable communication. Extensions to interacting RL, function approximation, or time-varying networks require further analysis. A promising future direction is to develop adaptive stepsize strategies that achieve communication savings without requiring prior knowledge of problem-dependent constants, while preserving strong convergence guarantees.

Funding and Competing Interests

This work received no external funding. The authors declare no competing interests.

References

[1] Dan Alistarh, Demjan Grubic, Jerry Li, Ryota Tomioka, and Milan Vojnovic. Qsgd: Communication-efficient sgd via gradient quantization and encoding. *Advances in neural information processing systems*, 30, 2017.

[2] Mario Arioli and Jennifer Scott. Chebyshev acceleration of iterative refinement. *Numerical Algorithms*, 66(3):591–608, 2014.

[3] Jalaj Bhandari, Daniel Russo, and Raghav Singal. A finite time analysis of temporal difference learning with linear function approximation. In *Conference on learning theory*, pages 1691–1692. PMLR, 2018.

[4] Pascal Bianchi, Gersende Fort, and Walid Hachem. Performance of a distributed stochastic approximation algorithm. *IEEE Transactions on Information Theory*, 59(11):7405–7418, 2013.

[5] Keith Bonawitz, Hubert Eichner, Wolfgang Grieskamp, Dzmitry Huba, Alex Ingerman, Vladimir Ivanov, Chloe Kiddon, Jakub Konečnỳ, Stefano Mazzocchi, Brendan McMahan, et al. Towards federated learning at scale: System design. *Proceedings of machine learning and systems*, 1:374–388, 2019.

[6] Léon Bottou, Frank E Curtis, and Jorge Nocedal. Optimization methods for large-scale machine learning. *SIAM review*, 60(2):223–311, 2018.

[7] Stephen Boyd, Persi Diaconis, and Lin Xiao. Fastest mixing markov chain on a graph. *SIAM Review*, 46(4):667–689, 2004.

[8] Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Gossip algorithms: Design, analysis and applications. In *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, volume 3, pages 1653–1664. IEEE, 2005.

[9] Tianyi Chen, Yuejiao Sun, Quan Xiao, and Wotao Yin. A single-timescale method for stochastic bilevel optimization. In *International Conference on Artificial Intelligence and Statistics*, pages 2466–2488. PMLR, 2022.

[10] Zaiwei Chen, Siva Theja Maguluri, Sanjay Shakkottai, and Karthikeyan Shanmugam. Finite-sample analysis of contractive stochastic approximation using smooth convex envelopes. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 8223–8234. Curran Associates, Inc., 2020.

[11] Zaiwei Chen, Siva Theja Maguluri, Sanjay Shakkottai, and Karthikeyan Shanmugam. A lyapunov theory for finite-sample guarantees of asynchronous q-learning and td-learning variants. *arXiv preprint arXiv:2102.01567*, 2021.

[12] Zaiwei Chen, Siva Theja Maguluri, and Martin Zubeldia. Concentration of contractive stochastic approximation: Additive and multiplicative noise. *arXiv preprint arXiv:2303.15740*, 2023.

[13] Thinh Doan, Siva Maguluri, and Justin Romberg. Finite-time analysis of distributed td (0) with linear function approximation on multi-agent reinforcement learning. In *International Conference on Machine Learning*, pages 1626–1635. PMLR, 2019.

[14] Thinh T Doan. Finite-time convergence rates of distributed local stochastic approximation. *Automatica*, 158:111294, 2023.

[15] Thinh T Doan, Siva Theja Maguluri, and Justin Romberg. Finite-time performance of distributed temporal-difference learning with linear function approximation. *SIAM Journal on Mathematics of Data Science*, 3(1):298–320, 2021.

[16] Drew Fudenberg and Jean Tirole. *Game theory*. MIT press, 1991.

[17] Peter Kairouz, H Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Kallista Bonawitz, Zachary Charles, Graham Cormode, Rachel Cummings, et al. Advances and open problems in federated learning. *Foundations and trends® in machine learning*, 14(1–2):1–210, 2021.

[18] Soummya Kar, José M. F. Moura, and Kavita Ramanan. Distributed parameter estimation in sensor networks: Nonlinear observation models and imperfect communication. *IEEE Transactions on Information Theory*, 58(6):3575–3605, 2012.

[19] Hassan K Khalil. *Nonlinear systems*. Prentice Hall, Upper Saddle River, NJ, 3 edition, 2002.

[20] Harold J Kushner and G George Yin. *Stochastic approximation and recursive algorithms and applications*. Springer, 2003.

[21] Yixuan Lin, Vijay Gupta, and Ji Liu. Finite-time error bounds for distributed linear stochastic approximation. *Automatica*, 159:111368, 2024.

[22] Yujun Lin, Song Han, Huizi Mao, Yu Wang, and Bill Dally. Deep gradient compression: Reducing the communication bandwidth for distributed training. In *International Conference on Learning Representations*, 2018.

[23] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[24] Guannan Qu and Adam Wierman. Finite-time analysis of asynchronous stochastic approximation and q-learning. In *Conference on Learning Theory*, pages 3185–3205. PMLR, 2020.

[25] Herbert Robbins and Sutton Monro. A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407, 1951.

[26] Kevin Scaman, Francis Bach, Sébastien Bubeck, Laurent Massoulié, and Yin Tat Lee. Optimal algorithms for non-smooth distributed optimization in networks. *Advances in Neural Information Processing Systems*, 31, 2018.

[27] Han Shen and Tianyi Chen. A single-timescale analysis for stochastic approximation with multiple coupled sequences. *Advances in Neural Information Processing Systems*, 35:17415–17429, 2022.

[28] Wei Shen, Minhui Huang, Jiawei Zhang, and Cong Shen. Stochastic smoothed gradient descent ascent for federated minimax optimization. In *International Conference on Artificial Intelligence and Statistics*, pages 3988–3996. PMLR, 2024.

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

[30] Sebastian U Stich. Local sgd converges fast and communicates little. In *International Conference on Learning Representations*, 2019.

[31] Haixiang Sun, Yang Xu, Jiefu Zhang, Xudong Wu, Zihan Zhou, Jun He, and Jiayu Chen. Core-Halo decomposition: Decentralizing large-scale fixed-point problems. *arXiv preprint arXiv:2605.08681*, 2026.

[32] Richard S Sutton, Andrew G Barto, et al. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

[33] Hoi-To Wai. On the convergence of consensus algorithms with markovian noise and gradient bias. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 4897–4902. IEEE, 2020.

[34] Pengfei Wang and Nenggan Zheng. Finite-time convergence rates of decentralized markovian stochastic approximation. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, pages 5082–5090, 2024.

[35] Lin Xiao, Stephen Boyd, and Sanjay Lall. Distributed average consensus with time-varying metropolis weights. Unpublished manuscript, June 2006, 2006.

[36] Yang Xu, Swetha Ganesh, and Vaneet Aggarwal. Efficient Q-Learning and actor–critic methods for robust average-reward reinforcement learning. In *Proceedings of the 42nd Conference on Uncertainty in Artificial Intelligence*, volume 337 of *Proceedings of Machine Learning Research*, pages 7613–7650. PMLR, 2026.

[37] Yang Xu, Washim Uddin Mondal, and Vaneet Aggarwal. Finite-sample analysis of policy evaluation for robust average reward reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

[38] Haishan Ye, Luo Luo, Ziang Zhou, and Tong Zhang. Multi-consensus decentralized accelerated gradient descent. *Journal of machine learning research*, 24(306):1–50, 2023.

[39] Haishan Ye, Wei Xiong, and Tong Zhang. Pmgt-vr: a decentralized proximal-gradient algorithmic framework with variance reduction. *IEEE transactions on pattern analysis and machine intelligence*, 2025.

[40] Honglin Yuan and Tengyu Ma. Federated accelerated stochastic gradient descent. *Advances in Neural Information Processing Systems*, 33:5332–5344, 2020.

[41] Sihan Zeng, Thinh T Doan, and Justin Romberg. Finite-time convergence rates of decentralized stochastic approximation with applications in multi-agent and multi-task learning. *IEEE Transactions on Automatic Control*, 68(5):2758–2773, 2023.

A Detailed Analysis of DLSA

This section completes the convergence analysis of DLSA. The algorithmic parameters are α, H and, for FastMix, n_ρ with prescribed η_w. The quantities $m, d, \beta, L, B, W, \rho, c_p, r_p$ describe the problem, and τ is the induced delay in communication rounds. The errors, Moreau envelope, auxiliary norms, and Lyapunov quantities are used only in the analysis.

A.1 Preliminaries

We collect the notation in Table 2 before introducing the Moreau envelope used in the proof.

Table 2: Notation used in the algorithm and analysis.

Quantity	Meaning
Algorithmic parameters	
α	Constant local-update stepsize.
H	Number of local updates in each communication round.
n_ρ	Number of gossip steps in one FastMix call.
η_w	FastMix momentum chosen from the network mixing parameter.
Problem quantities	
m, d	Number of agents and dimension of each local iterate.
$\bar{F}_i, \bar{F}, x^\star$	Local mean operators, their average, and its fixed point.
$\|\cdot\|_c, \beta$	General norm and contraction factor in Assumption 1.
L, B	Sampled-operator Lipschitz and fixed-point bounds in Assumption 1.
W, ρ	Gossip matrix and $\rho = \|W - \mathbf{1}\mathbf{1}^\top/m\|_2$.
W_b, ρ_b	Positive-semidefinite base used by FastMix and its disagreement factor.
c_p, r_p	Constants in the geometric Markov mixing bound.
$\tilde{\tau}_\epsilon, \tau$	Mixing horizon in sample transitions and the corresponding delay in communication rounds.
Quantities used in the proof	
R_k, S_k	Estimation error of the average iterate and consensus error at round k.
$R_{k,t}, S_{k,t}$	Corresponding errors after inner step t.
ρ_{eff}	Mixing factor of one communication phase.
$M, \theta, \|\cdot\|_s$	Moreau envelope, its parameter, and the auxiliary norm used in its construction.
$\|\cdot\|_M$	Norm satisfying $M(x) = \frac{1}{2}\|x\|_M^2$.
ℓ_{2c}, u_{2c}	Equivalence constants between the Euclidean norm and $\|\cdot\|_c$.
$\ell_{cm}, u_{cm}, \ell_{cs}$	Norm-equivalence constants in the Moreau-envelope construction.
L_{ms}, κ	Envelope smoothness constant and positive contraction quantity defined below.
$\Phi_{k,t}$	Envelope potential $M(\tilde{z}_t^k - x^\star)$.
$\Gamma_{k,\tau}, V^k$	Bound used for delayed iterates and the combined Lyapunov function defined in the analysis.
$c_\alpha, c_\rho, c_0, C_R, C_S$	Constants in the admissibility conditions and error bounds.

A.1.1 Moreau-envelope preliminaries

We introduce the generalized Moreau envelope $M(\cdot) : \mathbb{R}^d \to \mathbb{R}$, with parameter $\theta > 0$ and auxiliary norm $\|\cdot\|_s$,

$$M(x) := \min_{u \in \mathbb{R}^d} \left\{ \frac{1}{2}\|u\|_c^2 + \frac{1}{2\theta}\|u - x\|_s^2 \right\}.$$

The norm $\|\cdot\|_s$ is used to construct a smooth potential for the proof, and $\|\cdot\|_M$ is induced by the Moreau envelope. There exists a norm $\|\cdot\|_M$, such that for any $x \in \mathbb{R}^d$,

$$M(x) = \frac{1}{2}\|x\|_M^2, \qquad \|\nabla M(x)\|_M^\star = \|x\|_M, \qquad \langle \nabla M(x), x\rangle = \|x\|_M^2 = 2M(x). \tag{21}$$

Then, there exist constants $\ell_{cm}, u_{cm}, \ell_{cs} > 0$, such that

$$\ell_{cm}\|x\|_M \le \|x\|_c \le u_{cm}\|x\|_M, \quad \ell_{cs}\|x\|_s \le \|x\|_c, \qquad \forall x \in \mathbb{R}^d.$$

Moreover, $M(\cdot)$ is differentiable and L_{ms}-smooth w.r.t. $\|\cdot\|_s$, i.e., for all $x, y \in \mathbb{R}^d$,

$$M(y) \le M(x) + \langle \nabla M(x), y - x\rangle + \frac{L_{ms}}{2}\|y - x\|_s^2. \tag{22}$$

We also use the equivalent gradient Lipschitz property

$$\|\nabla M(x) - \nabla M(y)\|_s^\star \le L_{ms}\|x - y\|_s, \qquad \forall x, y \in \mathbb{R}^d. \tag{23}$$

In terms of the selection of the auxiliary norm $\|.\|_s$ given the β-contraction of $\bar{F}$ w.r.t. $\|.\|_c$, we assume the auxiliary norm $\|.\|_s$ is chosen well, such that

$$\kappa := 1 - \beta\frac{u_{cm}}{\ell_{cm}} \in (0, 1).$$

Markov-chain indexing and filtrations. For each agent i, the samples $\{Y_{i,t}^k\}_{k\ge 0, 0\le t\le H-1}$ form a single trajectory of a time-homogeneous Markov chain (with stationary distribution μ_i as assumed in the main text). It is sometimes convenient to flatten the double index via the global time $n := kH + t$ and write $Y_{i,t}^k = \widetilde{Y}_i^n$. We denote by $\mathcal{F}_{k,t}$ the sigma-field generated by all iterates and samples up to the beginning of inner step (k, t), i.e.,

$$\mathcal{F}_{k,t} := \sigma\Big(\{z_{i,s}^\ell, Y_{i,s}^\ell : 0 \le \ell < k, 0 \le s \le H-1, i \in [m]\} \cup \{z_{i,s}^k, Y_{i,s}^k : 0 \le s < t, i \in [m]\}\Big),$$

so that $z_{i,t}^k$ is $\mathcal{F}_{k,t}$-measurable and $Y_{i,t}^k$ is the next Markov state revealed at step (k, t).

A.2 Consensus error for Algorithm 1

Denote $J := \frac{1}{m}\mathbf{1}\mathbf{1}^\top$, then $\rho = \|W - J\|_2 \in [0, 1)$. The communication step can be written compactly as $X^{k+1} = (W \otimes I_d)Z_H^k$, and the post-communication disagreement satisfies $X^{k+1} - (\mathbf{1} \otimes \bar{x}^{k+1}) = ((W - J) \otimes I_d)\left(Z_H^k - (\mathbf{1} \otimes \bar{z}_H^k)\right)$. Although the analysis of consensus error is for one step of gossip communication with shrinking factor ρ, it can be directly replaced with ρ_{fm} if FastMix is adopted when the network is poorly connected, and the analysis keeps nearly unchanged.

This is the starting point for deriving contraction of consensus error.

Lemma 1 (Inner-loop drift bounds)**.** *Assume $0 < \alpha \le 1$. Fix an outer round $k \ge 0$ and consider the inner iterates $\{z_{i,t}^k\}_{t=0}^H$ generated by Algorithm 1. Then for every $t = 0, 1, \ldots, H-1$, the following inequalities hold (pathwise, i.e., for any realization of the Markov chains):*

$$R_{k,t+1} \le (1 - \alpha + \alpha L)\, R_{k,t} + \alpha L\, S_{k,t} + \alpha B, \tag{24}$$

$$S_{k,t+1} \le (1 - \alpha + 2\alpha L)\, S_{k,t} + 2\alpha\big(LR_{k,t} + B\big). \tag{25}$$

Moreover, letting

$$b := 1 - \alpha + \alpha L, \qquad a := 1 - \alpha + 2\alpha L,$$

we have the explicit unrolled forms for all $t = 1, 2, \ldots, H$:

$$R_{k,t} \le b^t R_{k,0} + \alpha L \sum_{s=0}^{t-1} b^{\,t-1-s} S_{k,s} + \alpha B \sum_{s=0}^{t-1} b^{\,t-1-s}, \tag{26}$$

$$S_{k,t} \le a^t S_{k,0} + 2\alpha \sum_{s=0}^{t-1} a^{\,t-1-s}\big(LR_{k,s} + B\big). \tag{27}$$

In particular, if $\alpha LH \le \frac{1}{4}$, then

$$S_{k,H} \le 2S_{k,0} + 4\alpha H\Big(L \max_{0\le s\le H-1} R_{k,s} + B\Big). \tag{28}$$

Proof. From the inner update $z_{i,t+1}^k = (1 - \alpha)z_{i,t}^k + \alpha F_i(z_{i,t}^k, Y_{i,t}^k)$ and the definition $\bar{z}_t^k = \frac{1}{m}\sum_{i=1}^m z_{i,t}^k$, we have

$$\bar{z}_{t+1}^k = (1 - \alpha)\bar{z}_t^k + \alpha\frac{1}{m}\sum_{i=1}^m F_i(z_{i,t}^k, Y_{i,t}^k). \tag{29}$$

Hence, by triangle inequality,

$$\begin{aligned}
R_{k,t+1} &= \|\bar{z}_{t+1}^k - x^\star\|_c \\
&= \Big\|(1-\alpha)(\bar{z}_t^k - x^\star) + \alpha\big(\frac{1}{m}\sum_{i=1}^m F_i(z_{i,t}^k, Y_{i,t}^k) - x^\star\big)\Big\|_c \\
&\le (1-\alpha)\|\bar{z}_t^k - x^\star\|_c + \alpha\Big\|\frac{1}{m}\sum_{i=1}^m F_i(z_{i,t}^k, Y_{i,t}^k) - x^\star\Big\|_c \\
&\le (1-\alpha)R_{k,t} + \alpha\frac{1}{m}\sum_{i=1}^m \|F_i(z_{i,t}^k, Y_{i,t}^k) - x^\star\|_c.
\end{aligned}$$

For each i, insert and subtract $F_i(x^\star, Y_{i,t}^k)$ and use Assumption 1(i),(iii):

$$\begin{aligned}
\|F_i(z_{i,t}^k, Y_{i,t}^k) - x^\star\|_c &\le \|F_i(z_{i,t}^k, Y_{i,t}^k) - F_i(x^\star, Y_{i,t}^k)\|_c + \|F_i(x^\star, Y_{i,t}^k) - x^\star\|_c \\
&\le L\|z_{i,t}^k - x^\star\|_c + B.
\end{aligned}$$

Also,

$$\|z_{i,t}^k - x^\star\|_c \le \|z_{i,t}^k - \bar{z}_t^k\|_c + \|\bar{z}_t^k - x^\star\|_c = \|z_{i,t}^k - \bar{z}_t^k\|_c + R_{k,t}.$$

Averaging over i gives

$$\frac{1}{m}\sum_{i=1}^m \|z_{i,t}^k - x^\star\|_c \le S_{k,t} + R_{k,t}.$$

Combining the last three displays yields

$$\begin{aligned}
R_{k,t+1} &\le (1-\alpha)R_{k,t} + \alpha\Big(L(S_{k,t} + R_{k,t}) + B\Big) \\
&= (1-\alpha+\alpha L)R_{k,t} + \alpha L S_{k,t} + \alpha B,
\end{aligned}$$

which is (24).

From the inner update and (29),

$$z_{i,t+1}^k - \bar{z}_{t+1}^k = (1-\alpha)(z_{i,t}^k - \bar{z}_t^k) + \alpha\Big(F_i(z_{i,t}^k, Y_{i,t}^k) - \frac{1}{m}\sum_{i=1}^m F_j(z_{j,t}^k, Y_{j,t}^k)\Big).$$

Taking $\|\cdot\|_c$ and applying triangle inequality,

$$\|z_{i,t+1}^k - \bar{z}_{t+1}^k\|_c \le (1-\alpha)\|z_{i,t}^k - \bar{z}_t^k\|_c + \alpha\Big\|F_i(z_{i,t}^k, Y_{i,t}^k) - \frac{1}{m}\sum_{i=1}^m F_j(z_{j,t}^k, Y_{j,t}^k)\Big\|_c. \tag{30}$$

We further split the last term by adding and subtracting $F_i(\bar{z}_t^k, Y_{i,t}^k)$ and $\frac{1}{m}\sum_{i=1}^m F_j(\bar{z}_t^k, Y_{j,t}^k)$:

$$\begin{aligned}
\Big\|F_i(z_{i,t}^k, Y_{i,t}^k) - \frac{1}{m}\sum_{i=1}^m F_j(z_{j,t}^k, Y_{j,t}^k)\Big\|_c &\le \underbrace{\|F_i(z_{i,t}^k, Y_{i,t}^k) - F_i(\bar{z}_t^k, Y_{i,t}^k)\|_c}_{\text{(A)}} \\
+ \underbrace{\Big\|F_i(\bar{z}_t^k, Y_{i,t}^k) - \frac{1}{m}\sum_{i=1}^m F_j(\bar{z}_t^k, Y_{j,t}^k)\Big\|_c}_{\text{(B)}} &+ \underbrace{\Big\|\frac{1}{m}\sum_{i=1}^m \big(F_j(\bar{z}_t^k, Y_{j,t}^k) - F_j(z_{j,t}^k, Y_{j,t}^k)\big)\Big\|_c}_{\text{(C)}}.
\end{aligned} \tag{31}$$

We bound (A), (B), (C) separately.

For (A), by Assumption 1 (i),

$$\text{(A)} \le L\|z_{i,t}^k - \bar{z}_t^k\|_c.$$

For (B), using $\|u - \frac{1}{m}\sum_{i=1}^m v_j\|_c \le \|u\|_c + \|\frac{1}{m}\sum_{i=1}^m v_j\|_c \le \|u\|_c + \frac{1}{m}\sum_{i=1}^m \|v_j\|_c$, we obtain

$$\text{(B)} \le \|F_i(\bar{z}_t^k, Y_{i,t}^k)\|_c + \frac{1}{m}\sum_{i=1}^m \|F_j(\bar{z}_t^k, Y_{j,t}^k)\|_c.$$

For any i, by Assumption 1 (i), (iii),

$$\begin{aligned}\|F_i(\bar{z}_t^k, Y_{i,t}^k)\|_c &\le \|F_i(\bar{z}_t^k, Y_{i,t}^k) - F_i(x^\star, Y_{i,t}^k)\|_c + \|F_i(x^\star, Y_{i,t}^k)\|_c \\ &\le L\|\bar{z}_t^k - x^\star\|_c + B = LR_{k,t} + B.\end{aligned}$$

Hence,

$$(\mathrm{B}) \le 2(LR_{k,t} + B).$$

For (C), by Jensen and Assumption 1 (i),

$$(\mathrm{C}) \le \frac{1}{m}\sum_{i=1}^m \|F_i(\bar{z}_t^k, Y_{i,t}^k) - F_i(z_{i,t}^k, Y_{i,t}^k)\|_c \le \frac{1}{m}\sum_{i=1}^m L\|\bar{z}_t^k - z_{i,t}^k\|_c = LS_{k,t}.$$

Plugging the bounds for (A)–(C) into (31) and then into (30) yields

$$\|z_{i,t+1}^k - \bar{z}_{t+1}^k\|_c \le (1-\alpha)\|z_{i,t}^k - \bar{z}_t^k\|_c + \alpha\Big(L\|z_{i,t}^k - \bar{z}_t^k\|_c + 2(LR_{k,t} + B) + LS_{k,t}\Big).$$

Rearranging,

$$\|z_{i,t+1}^k - \bar{z}_{t+1}^k\|_c \le (1 - \alpha + \alpha L)\|z_{i,t}^k - \bar{z}_t^k\|_c + \alpha L S_{k,t} + 2\alpha(LR_{k,t} + B).$$

Averaging this inequality over $i \in [m]$ gives

$$\begin{aligned}S_{k,t+1} &= \frac{1}{m}\sum_{i=1}^m \|z_{i,t+1}^k - \bar{z}_{t+1}^k\|_c \\ &\le (1 - \alpha + \alpha L)\frac{1}{m}\sum_{i=1}^m \|z_{i,t}^k - \bar{z}_t^k\|_c + \alpha L S_{k,t} + 2\alpha(LR_{k,t} + B) \\ &= (1 - \alpha + 2\alpha L)S_{k,t} + 2\alpha(LR_{k,t} + B),\end{aligned}$$

which is (25).

We first show (27). Starting from (25), for $t \ge 0$,

$$S_{k,t+1} \le aS_{k,t} + 2\alpha(LR_{k,t} + B).$$

Apply this inequality recursively:

$$\begin{aligned}&S_{k,1} \le aS_{k,0} + 2\alpha(LR_{k,0} + B), \\ &S_{k,2} \le aS_{k,1} + 2\alpha(LR_{k,1} + B) \le a^2 S_{k,0} + 2\alpha\big(a(LR_{k,0} + B) + (LR_{k,1} + B)\big),\end{aligned}$$

and by induction one obtains

$$S_{k,t} \le a^t S_{k,0} + 2\alpha\sum_{s=0}^{t-1} a^{\,t-1-s}(LR_{k,s} + B),$$

which is (27). The bound (28) follows by using $0 \le a \le 1 + 2\alpha L$ and, for $t = H$,

$$\sum_{s=0}^{H-1} a^{H-1-s} \le H\max\{1, a^H\} \le H(1 + 2\alpha L)^H \le 2H,$$

together with $\max_{0\le s\le H-1}(LR_{k,s} + B)$ pulled out of the sum in (27).

For every $t \ge 0$ we have

$$R_{k,t+1} \le bR_{k,t} + \alpha L S_{k,t} + \alpha B, \qquad \text{where } b := 1 - \alpha + \alpha L. \tag{32}$$

We now unroll (32) explicitly. First,

$$R_{k,1} \le bR_{k,0} + \alpha L S_{k,0} + \alpha B.$$

Applying (32) again with $t = 1$ yields

$$\begin{aligned} R_{k,2} &\le bR_{k,1} + \alpha L S_{k,1} + \alpha B \\ &\le b\Big(bR_{k,0} + \alpha L S_{k,0} + \alpha B\Big) + \alpha L S_{k,1} + \alpha B \\ &= b^2 R_{k,0} + \alpha L\Big(bS_{k,0} + S_{k,1}\Big) + \alpha B\Big(b+1\Big). \end{aligned}$$

Similarly, applying (32) once more gives

$$\begin{aligned} R_{k,3} &\le bR_{k,2} + \alpha L S_{k,2} + \alpha B \\ &\le b\Big(b^2 R_{k,0} + \alpha L(bS_{k,0} + S_{k,1}) + \alpha B(b+1)\Big) + \alpha L S_{k,2} + \alpha B \\ &= b^3 R_{k,0} + \alpha L\Big(b^2 S_{k,0} + bS_{k,1} + S_{k,2}\Big) + \alpha B\Big(b^2 + b + 1\Big). \end{aligned}$$

From these first few cases the pattern is clear. We prove by induction that for any integer $t \ge 1$

$$R_{k,t} \le b^t R_{k,0} + \alpha L \sum_{s=0}^{t-1} b^{\,t-1-s} S_{k,s} + \alpha B \sum_{s=0}^{t-1} b^{\,t-1-s}. \tag{33}$$

For the base case $t = 1$, (33) is exactly $R_{k,1} \le bR_{k,0} + \alpha L S_{k,0} + \alpha B$. Assume (33) holds for some $t \ge 1$. Then by (32),

$$\begin{aligned} R_{k,t+1} &\le bR_{k,t} + \alpha L S_{k,t} + \alpha B \\ &\le b\Big(b^t R_{k,0} + \alpha L \sum_{s=0}^{t-1} b^{\,t-1-s} S_{k,s} + \alpha B \sum_{s=0}^{t-1} b^{\,t-1-s}\Big) + \alpha L S_{k,t} + \alpha B \\ &= b^{t+1} R_{k,0} + \alpha L \sum_{s=0}^{t-1} b^{\,t-s} S_{k,s} + \alpha B \sum_{s=0}^{t-1} b^{\,t-s} + \alpha L S_{k,t} + \alpha B \\ &= b^{t+1} R_{k,0} + \alpha L \sum_{s=0}^{t} b^{\,t-s} S_{k,s} + \alpha B \sum_{s=0}^{t} b^{\,t-s}, \end{aligned}$$

which is exactly (33) with t replaced by $t+1$. This completes the induction and proves (26). □

Lemma 2 (Consensus error after one communication)**.** *Fix $k \ge 0$ and let $\{z_{i,H}^k\}_{i=1}^m$ be the iterates of H local steps from Algorithm 1, and let $\{x_i^{k+1}\}_{i=1}^m$ be the iterates after local communications,*

$$x_i^{k+1} = \sum_{j=1}^m w_{ij} z_{j,H}^k, \qquad i \in [m].$$

Then the one-communication step satisfies

$$S_{k+1} \le \rho\, \frac{u_{2c}}{\ell_{2c}}\, \sqrt{m}\, S_{k,H}, \tag{34}$$

where $\ell_{2c}, u_{2c} > 0$ are norm equivalence parameters, such that

$$\ell_{2c}\|v\|_2 \le \|v\|_c \le u_{2c}\|v\|_2, \quad \forall v \in \mathbb{R}^d. \tag{35}$$

In particular, if $\alpha LH \le \frac{1}{4}$, then Lemma 1 implies

$$S_{k+1} \le \rho\, \frac{u_{2c}}{\ell_{2c}}\, \sqrt{m}\, \Big(2S_k + 4\alpha H\big(L \max_{0 \le s \le H-1} R_{k,s} + B\big)\Big). \tag{36}$$

Proof. By the property of mixing matrix W, we have

$$\left(\frac{1}{m}\sum_{i=1}^m \|x_i^{k+1} - \bar{x}^{k+1}\|_2^2\right)^{1/2} \le \rho \left(\frac{1}{m}\sum_{i=1}^m \|z_{i,H}^k - \bar{z}_H^k\|_2^2\right)^{1/2}. \tag{37}$$

We bound S_{k+1} from above using (35),

$$S_{k+1} = \frac{1}{m}\sum_{i=1}^{m} \|x_i^{k+1} - \bar{x}^{k+1}\|_c \le \frac{u_{2c}}{m}\sum_{i=1}^{m} \|x_i^{k+1} - \bar{x}^{k+1}\|_2 \\ \le u_{2c}\left(\frac{1}{m}\sum_{i=1}^{m} \|x_i^{k+1} - \bar{x}^{k+1}\|_2^2\right)^{1/2}. \tag{38}$$

On the other hand,

$$\left(\frac{1}{m}\sum_{i=1}^{m} \|z_{i,H}^k - \bar{z}_H^k\|_2^2\right)^{1/2} \le \frac{1}{\ell_{2c}}\left(\frac{1}{m}\sum_{i=1}^{m} \|z_{i,H}^k - \bar{z}_H^k\|_c^2\right)^{1/2} \le \frac{\sqrt{m}}{\ell_{2c}} S_{k,H}. \tag{39}$$

Then

$$S_{k+1} \le u_{2c}\left(\frac{1}{m}\sum_{i=1}^{m} \|x_i^{k+1} - \bar{x}^{k+1}\|_2^2\right)^{1/2} \\ \le u_{2c}\rho\left(\frac{1}{m}\sum_{i=1}^{m} \|z_{i,H}^k - \bar{z}_H^k\|_2^2\right)^{1/2} \\ \le \rho\,\frac{u_{2c}}{\ell_{2c}}\sqrt{m}\,S_{k,H}.$$

If $\alpha LH \le \frac{1}{4}$, Lemma 1 gives $S_{k,H} \le 2S_k + 4\alpha H(L\max_{0\le s\le H-1} R_{k,s} + B)$. Substituting this into (34) yields (36). □

A.3 Estimation gap along the inner-loop average

Define the potential

$$\Phi_{k,t} := M(\bar{z}_t^k - x^\star), \qquad t = 0, 1, \dots, H.$$

Then $\Phi_{k,0} = M(\bar{x}^k - x^\star)$ and $\Phi_{k,H} = M(\bar{x}^{k+1} - x^\star)$.

Denote $\Delta_{k,t} := \bar{z}_{t+1}^k - \bar{z}_t^k$, then by local update

$$z_{i,t+1}^k = z_{i,t}^k + \alpha(F_i(z_{i,t}^k, Y_{i,t}^k) - z_{i,t}^k)$$

we have

$$\Delta_{k,t} = \alpha\left(\frac{1}{m}\sum_{i=1}^{m} F_i(z_{i,t}^k, Y_{i,t}^k) - \bar{z}_t^k\right) = \alpha\big(\bar{F}(\bar{z}_t^k) - \bar{z}_t^k\big) + \alpha\,\mathsf{E}_{k,t}^{\mathrm{cons}} + \alpha\,\mathsf{E}_{k,t}^{\mathrm{mk}}, \tag{40}$$

where

$$\mathsf{E}_{k,t}^{\mathrm{cons}} := \frac{1}{m}\sum_{i=1}^{m}\Big(F_i(z_{i,t}^k, Y_{i,t}^k) - F_i(\bar{z}_t^k, Y_{i,t}^k)\Big), \quad \mathsf{E}_{k,t}^{\mathrm{mk}} := \frac{1}{m}\sum_{i=1}^{m}\Big(F_i(\bar{z}_t^k, Y_{i,t}^k) - \bar{F}_i(\bar{z}_t^k)\Big). \tag{41}$$

In addition, we denote $\mathcal{T}_{k,t}^{\mathrm{mk}} := \alpha\,\mathbb{E}[\langle \nabla M(\bar{z}_t^k - x^\star), \mathsf{E}_{k,t}^{\mathrm{mk}}\rangle]$, a quantity relevant to Markovian noise.

Lemma 3 (Inner-step descent inequality)**.** *Fix $k \ge 0$ and $t \in \{0, 1, \dots, H-1\}$, then*

$$\mathbb{E}[\Phi_{k,t+1}] \le \Big(1 - \alpha\kappa + \frac{3L_{ms}\alpha^2(L+1)^2u_{cm}^2}{\ell_{cs}^2}\Big)\,\mathbb{E}[\Phi_{k,t}] + \mathcal{T}_{k,t}^{\mathrm{mk}} \\ + \Big(\frac{\alpha L^2 u_{cm}^2}{2\kappa\,\ell_{cm}^4} + \frac{3L_{ms}\alpha^2L^2}{2\ell_{cs}^2}\Big)\,\mathbb{E}[S_{k,t}^2] + \frac{3L_{ms}\alpha^2}{2\ell_{cs}^2}\,B^2. \tag{42}$$

Proof. Applying (22) with $x = \bar{z}_t^k - x^\star$ and $y = \bar{z}_{t+1}^k - x^\star$ gives

$$\Phi_{k,t+1} \le \Phi_{k,t} + \big\langle \nabla M(\bar{z}_t^k - x^\star), \Delta_{k,t}\big\rangle + \frac{L_{ms}}{2}\|\Delta_{k,t}\|_s^2. \tag{43}$$

For $\langle \nabla M(\bar{z}_t^k - x^\star), \Delta_{k,t} \rangle$, we separate it into three parts using (40),

$$\begin{aligned}&\langle \nabla M(\bar{z}_t^k - x^\star), \Delta_{k,t} \rangle \\ &= \langle \nabla M(\bar{z}_t^k - x^\star), \alpha(\bar{F}(\bar{z}_t^k) - \bar{z}_t^k) \rangle + \langle \nabla M(\bar{z}_t^k - x^\star), \alpha \mathsf{E}_{k,t}^{\mathrm{cons}} \rangle + \langle \nabla M(\bar{z}_t^k - x^\star), \alpha \mathsf{E}_{k,t}^{\mathrm{mk}} \rangle.\end{aligned}$$

Then, we need to bound the terms $\langle \nabla M(\bar{z}_t^k - x^\star), \alpha(\bar{F}(\bar{z}_t^k) - \bar{z}_t^k) \rangle$, $\langle \nabla M(\bar{z}_t^k - x^\star), \alpha \mathsf{E}_{k,t}^{\mathrm{cons}} \rangle$ and $\|\Delta_{k,t}\|_s^2$. The term $\langle \nabla M(\bar{z}_t^k - x^\star), \alpha \mathsf{E}_{k,t}^{\mathrm{mk}} \rangle$ is exactly $\mathcal{T}_{k,t}^{\mathrm{mk}}$ after taking expectation.

Bound $\langle \nabla M(\bar{z}_t^k - x^\star), \alpha(\bar{F}(\bar{z}_t^k) - \bar{z}_t^k) \rangle$:

Let $x_{k,t} = \bar{z}_t^k - x^\star$ as in the statement. Using $\bar{F}(x^\star) = x^\star$,

$$\bar{F}(\bar{z}_t^k) - \bar{z}_t^k = \big(\bar{F}(\bar{z}_t^k) - \bar{F}(x^\star)\big) - (\bar{z}_t^k - x^\star).$$

Therefore, by (21),

$$\begin{aligned}\langle \nabla M(\bar{z}_t^k - x^\star), \bar{F}(\bar{z}_t^k) - \bar{z}_t^k \rangle &= -\langle \nabla M(\bar{z}_t^k - x^\star), \bar{z}_t^k - x^\star \rangle + \langle \nabla M(\bar{z}_t^k - x^\star), \bar{F}(\bar{z}_t^k) - \bar{F}(x^\star) \rangle \\ &= -2M(\bar{z}_t^k - x^\star) + \langle \nabla M(\bar{z}_t^k - x^\star), \bar{F}(\bar{z}_t^k) - x^\star \rangle. \end{aligned} \tag{44}$$

Bound the second term in (44) using

$$\langle \nabla M(\bar{z}_t^k - x^\star), \bar{F}(\bar{z}_t^k) - x^\star \rangle \le \|\nabla M(\bar{z}_t^k - x^\star)\|_M^\star \|\bar{F}(\bar{z}_t^k) - x^\star\|_M.$$

By (21), $\|\nabla M(\bar{z}_t^k - x^\star)\|_M^\star = \|\bar{z}_t^k - x^\star\|_M$. Next, by β-contraction of $\bar{F}$ in $\|\cdot\|_c$,

$$\|\bar{F}(\bar{z}_t^k) - x^\star\|_M \le \frac{1}{\ell_{cm}} \|\bar{F}(\bar{z}_t^k) - x^\star\|_c \le \frac{\beta}{\ell_{cm}} \|\bar{z}_t^k - x^\star\|_c \le \frac{\beta u_{cm}}{\ell_{cm}} \|\bar{z}_t^k - x^\star\|_M.$$

Combining the last two displays yields

$$\langle \nabla M(\bar{z}_t^k - x^\star), \bar{F}(\bar{z}_t^k) - x^\star \rangle \le \frac{\beta u_{cm}}{\ell_{cm}} \|\bar{z}_t^k - x^\star\|_M^2 = \frac{2\beta u_{cm}}{\ell_{cm}} M(\bar{z}_t^k - x^\star).$$

Plugging this into (44) gives

$$\langle \nabla M(\bar{z}_t^k - x^\star), \bar{F}(\bar{z}_t^k) - \bar{z}_t^k \rangle \le -2\Big(1 - \beta \frac{u_{cm}}{\ell_{cm}}\Big) M(\bar{z}_t^k - x^\star) = -2\kappa\, \Phi_{k,t}. \tag{45}$$

Bound $\langle \nabla M(\bar{z}_t^k - x^\star), \alpha \mathsf{E}_{k,t}^{\mathrm{cons}} \rangle$:

Since

$$\|\nabla M(\bar{z}_t^k - x^\star)\|_M^\star = \|\bar{z}_t^k - x^\star\|_M \le \frac{1}{\ell_{cm}} \|\bar{z}_t^k - x^\star\|_c = \frac{1}{\ell_{cm}} R_{k,t}. \tag{46}$$

$$\|\mathsf{E}_{k,t}^{\mathrm{cons}}\|_M \le \frac{1}{\ell_{cm}} \|\mathsf{E}_{k,t}^{\mathrm{cons}}\|_c \le \frac{1}{\ell_{cm}} \cdot \frac{1}{m} \sum_{i=1}^m \|F_i(z_{i,t}^k, Y_{i,t}^k) - F_i(\bar{z}_t^k, Y_{i,t}^k)\|_c \le \frac{L}{\ell_{cm}} S_{k,t},$$

we have

$$\langle \nabla M(\bar{z}_t^k - x^\star), \mathsf{E}_{k,t}^{\mathrm{cons}} \rangle \le \|\nabla M(\bar{z}_t^k - x^\star)\|_M^\star \|\mathsf{E}_{k,t}^{\mathrm{cons}}\|_M \le \frac{L}{\ell_{cm}^2} R_{k,t} S_{k,t}. \tag{47}$$

We now absorb the resulting $\Phi_{k,t}$ contribution into the contractive term (45). We have $R_{k,t} = \|\bar{z}_t^k - x^\star\|_c \le u_{cm} \|\bar{z}_t^k - x^\star\|_M = u_{cm} \sqrt{2\Phi_{k,t}}$. Thus

$$\frac{L}{\ell_{cm}^2} R_{k,t} S_{k,t} \le \frac{L u_{cm}}{\ell_{cm}^2} \sqrt{2\Phi_{k,t}}\, S_{k,t} \le \kappa\, \Phi_{k,t} + \frac{L^2 u_{cm}^2}{2\kappa\, \ell_{cm}^4} S_{k,t}^2.$$

Therefore,

$$\alpha \langle \nabla M(\bar{z}_t^k - x^\star), \mathsf{E}_{k,t}^{\mathrm{cons}} \rangle \le \alpha\kappa\, \Phi_{k,t} + \frac{\alpha L^2 u_{cm}^2}{2\kappa\, \ell_{cm}^4} S_{k,t}^2. \tag{48}$$

Bound $\|\Delta_{k,t}\|_s^2$:

Using (40),

$$\|\Delta_{k,t}\|_s^2 \le \frac{\alpha^2}{\ell_{cs}^2}\left\|\frac{1}{m}\sum_{i=1}^m F_i(z_{i,t}^k, Y_{i,t}^k) - \bar{z}_t^k\right\|_c^2. \tag{49}$$

Insert and subtract $F_i(x^\star, Y_{i,t}^k)$ and $x^\star$:

$$\begin{aligned}
&\frac{1}{m}\sum_{i=1}^m F_i(z_{i,t}^k, Y_{i,t}^k) - \bar{z}_t^k \\
&= \frac{1}{m}\sum_{i=1}^m \big(F_i(z_{i,t}^k, Y_{i,t}^k) - F_i(x^\star, Y_{i,t}^k)\big) \\
&\quad + \frac{1}{m}\sum_{i=1}^m \big(F_i(x^\star, Y_{i,t}^k) - x^\star\big) + (x^\star - \bar{z}_t^k).
\end{aligned}$$

By triangle inequality and Assumption 1(i),(iii),

$$\begin{aligned}
&\|\frac{1}{m}\sum_{i=1}^m F_i(z_{i,t}^k, Y_{i,t}^k) - \bar{z}_t^k\|_c \le \frac{1}{m}\sum_{i=1}^m L\|z_{i,t}^k - x^\star\|_c + B + \|\bar{z}_t^k - x^\star\|_c \\
&\le L\Big(\frac{1}{m}\sum_{i=1}^m \|z_{i,t}^k - \bar{z}_t^k\|_c + \|\bar{z}_t^k - x^\star\|_c\Big) + B + R_{k,t} = (L+1)R_{k,t} + LS_{k,t} + B.
\end{aligned}$$

Then

$$\|\frac{1}{m}\sum_{i=1}^m F_i(z_{i,t}^k, Y_{i,t}^k) - \bar{z}_t^k\|_c^2 \le 3\Big((L+1)^2 R_{k,t}^2 + L^2 S_{k,t}^2 + B^2\Big).$$

Plug into (49) and use $R_{k,t}^2 = \|\bar{z}_t^k - x^\star\|_c^2 \le u_{cm}^2\|\bar{z}_t^k - x^\star\|_M^2 = 2u_{cm}^2\Phi_{k,t}$,

$$\frac{L_{ms}}{2}\|\Delta_{k,t}\|_s^2 \le \frac{3L_{ms}\alpha^2(L+1)^2u_{cm}^2}{\ell_{cs}^2}\,\Phi_{k,t} + \frac{3L_{ms}\alpha^2L^2}{2\ell_{cs}^2}\,S_{k,t}^2 + \frac{3L_{ms}\alpha^2}{2\ell_{cs}^2}\,B^2. \tag{50}$$

Combine (45), (48), (50) and take expectation, then yield exactly (42). □

We now bound $\mathcal{T}_{k,t}^{\mathrm{mk}}$ under the geometric mixing assumption specified in (4). A delay of τ communication rounds provides at least $H\tau$ Markov transitions before the current sample. For a constant stepsize $\alpha > 0$, choose the mixing tolerance and the integer delay as

$$\begin{aligned}
\epsilon(\alpha) &:= LL_{ms}\frac{\ell_{cs}^2}{\ell_{cm}^2}\alpha, \\
\tau &:= \max\left\{1, \left\lceil \frac{1}{H(1-r_p)}\log\frac{c_p}{\epsilon(\alpha)}\right\rceil\right\}.
\end{aligned} \tag{51}$$

We take α sufficiently small that $\epsilon(\alpha) \in (0,1)$. Since $r_p^s \le \exp(-(1-r_p)s)$, this choice gives $c_p r_p^{H\tau} \le \epsilon(\alpha)$ and hence $H\tau \ge \tilde{\tau}_{\epsilon(\alpha)}$. The maximum with 1 ensures a positive integer delay even when H exceeds the mixing horizon. It also gives

$$\tau \le 1 + \frac{\max\{0, \log(c_p/\epsilon(\alpha))\}}{H(1-r_p)}.$$

Thus reducing α tightens the mixing tolerance and increases the required horizon logarithmically, while increasing H reduces its conversion to communication rounds until the lower bound $\tau = 1$ is reached. The delay is fixed for fixed α and H, and the stepsize and network conditions below are checked with this value of τ. The additive 1 in the main theorem retains the effect of integer rounding. For any $k \ge \tau$, $t \in \{0, \dots, H-1\}$, $i \in [m]$, the chain has evolved for at least $H\tau$ steps since $\mathcal{F}_{k-\tau,0}$, by (4),

$$d_{\mathrm{TV}}\big(\mathcal{L}(Y_{i,t}^k \,|\, \mathcal{F}_{k-\tau,0}), \mu_i\big) \le c_p r_p^{H\tau} \le \epsilon. \tag{52}$$

We now perform a four-term decomposition as follows:

$$\begin{aligned}\mathcal{T}_{k,t}^{\mathrm{mk}} = \frac{\alpha}{m}\sum_{i=1}^{m}\mathbb{E}[\Big\langle \nabla M(\bar{z}_t^k - x^\star) - \nabla M(\bar{x}^{k-\tau} - x^\star),\, F_i(\bar{z}_t^k, Y_{i,t}^k) - \bar{F}_i(\bar{z}_t^k)\Big\rangle] \\ + \frac{\alpha}{m}\sum_{i=1}^{m}\mathbb{E}[\Big\langle \nabla M(\bar{x}^{k-\tau} - x^\star),\, F_i(\bar{x}^{k-\tau}, Y_{i,t}^k) - \bar{F}_i(\bar{x}^{k-\tau})\Big\rangle] \\ + \frac{\alpha}{m}\sum_{i=1}^{m}\mathbb{E}[\Big\langle \nabla M(\bar{x}^{k-\tau} - x^\star),\, F_i(\bar{z}_t^k, Y_{i,t}^k) - F_i(\bar{x}^{k-\tau}, Y_{i,t}^k)\Big\rangle] \\ + \frac{\alpha}{m}\sum_{i=1}^{m}\mathbb{E}[\Big\langle \nabla M(\bar{x}^{k-\tau} - x^\star),\, \bar{F}_i(\bar{x}^{k-\tau}) - \bar{F}_i(\bar{z}_t^k)\Big\rangle] \end{aligned} \tag{53}$$

gives the corresponding decomposition of $\mathcal{T}_{k,t}^{\mathrm{mk}}$ into four parts.

$$\begin{aligned}&\Big|\Big\langle \nabla M(\bar{z}_t^k - x^\star) - \nabla M(\bar{x}^{k-\tau} - x^\star),\, F_i(\bar{z}_t^k, Y_{i,t}^k) - \bar{F}_i(\bar{z}_t^k)\Big\rangle\Big| \\ &\le \|\nabla M(\bar{z}_t^k - x^\star) - \nabla M(\bar{x}^{k-\tau} - x^\star)\|_s^\star \|F_i(\bar{z}_t^k, Y_{i,t}^k) - \bar{F}_i(\bar{z}_t^k)\|_s \\ &\le L_{ms}\|\bar{z}_t^k - \bar{x}^{k-\tau}\|_s \cdot \frac{1}{\ell_{cs}}\|F_i(\bar{z}_t^k, Y_{i,t}^k) - \bar{F}_i(\bar{z}_t^k)\|_c \\ &\le \frac{L_{ms}}{\ell_{cs}^2}\,\|\bar{z}_t^k - \bar{x}^{k-\tau}\|_c \cdot \|F_i(\bar{z}_t^k, Y_{i,t}^k) - \bar{F}_i(\bar{z}_t^k)\|_c \qquad (54) \\ &\le \frac{2L_{ms}}{\ell_{cs}^2}\,\|\bar{z}_t^k - \bar{x}^{k-\tau}\|_c\,(LR_{k,t} + B), \qquad (55)\end{aligned}$$

where the last inequality is due to

$$\begin{aligned}&\|F_i(\bar{z}_t^k, Y_{i,t}^k) - \bar{F}_i(\bar{z}_t^k)\|_c \le \big\|F_i(\bar{z}_t^k, Y_{i,t}^k)\big\|_c + \big\|\bar{F}_i(\bar{z}_t^k)\big\|_c \\ &\le \|F_i(\bar{z}_t^k, Y_{i,t}^k) - F_i(x^\star, Y_{i,t}^k)\|_c + \|F_i(x^\star, Y_{i,t}^k)\|_c + \big\|\bar{F}_i(\bar{z}_t^k) - \bar{F}_i(x^\star)\big\|_c + \big\|\bar{F}_i(x^\star)\big\|_c \\ &\le 2(LR_{k,t} + B).\end{aligned}$$

Averaging over i yields

$$\begin{aligned}&\frac{\alpha}{m}\sum_{i=1}^{m}\mathbb{E}\Big[\Big\langle \nabla M(\bar{z}_t^k - x^\star) - \nabla M(\bar{x}^{k-\tau} - x^\star),\, F_i(\bar{z}_t^k, Y_{i,t}^k) - \bar{F}_i(\bar{z}_t^k)\Big\rangle\Big] \\ &\le \frac{2\alpha L_{ms}}{\ell_{cs}^2}\,\mathbb{E}[\|\bar{z}_t^k - \bar{x}^{k-\tau}\|_c\,(LR_{k,t} + B)].\end{aligned} \tag{56}$$

$$\begin{aligned}&\Big|\mathbb{E}[\Big\langle \nabla M(\bar{x}^{k-\tau} - x^\star),\, F_i(\bar{x}^{k-\tau}, Y_{i,t}^k) - \bar{F}_i(\bar{x}^{k-\tau})\Big\rangle\Big|\,\mathcal{F}_{k-\tau,0}]\Big| \\ &\le \|\nabla M(\bar{x}^{k-\tau} - x^\star)\|_M^\star \Big\|\mathbb{E}[F_i(\bar{x}^{k-\tau}, Y_{i,t}^k) - \bar{F}_i(\bar{x}^{k-\tau})\Big|\,\mathcal{F}_{k-\tau,0}]\Big\|_M \\ &\le \frac{1}{\ell_{cm}^2}\,\|\bar{x}^{k-\tau} - x^\star\|_c \Big\|\mathbb{E}[F_i(\bar{x}^{k-\tau}, Y_{i,t}^k)\Big|\,\mathcal{F}_{k-\tau,0}] - \bar{F}_i(\bar{x}^{k-\tau})\Big\|_c. \qquad (57)\end{aligned}$$

By the TV bound (52) and the standard inequality $\|\sum_y (p(y) - \mu(y))g(y)\|_c \le 2\, d_{\mathrm{TV}}(p,\mu)\, \sup_y \|g(y)\|_c$, we have

$$\Big\|\mathbb{E}[F_i(\bar{x}^{k-\tau}, Y_{i,t}^k)\Big|\,\mathcal{F}_{k-\tau,0}] - \bar{F}_i(\bar{x}^{k-\tau})\Big\|_c \le 2\epsilon \sup_{y\in\mathcal{Y}}\|F_i(\bar{x}^{k-\tau}, y)\|_c \le 2\epsilon\big(L\|\bar{x}^{k-\tau} - x^\star\|_c + B\big). \tag{58}$$

Plugging (58) into (57), taking expectation, and averaging over i yields

$$\begin{aligned}&\frac{\alpha}{m}\sum_{i=1}^{m}\mathbb{E}\Big[\Big\langle\nabla M(\bar{x}^{k-\tau}-x^\star),\,F_i(\bar{x}^{k-\tau},Y_{i,t}^k)-\bar{F}_i(\bar{x}^{k-\tau})\Big\rangle\Big]\\&\le\frac{2\alpha\epsilon}{\ell_{cm}^2}\,\mathbb{E}[\|\bar{x}^{k-\tau}-x^\star\|_c\big(L\|\bar{x}^{k-\tau}-x^\star\|_c+B\big)].\end{aligned}\tag{59}$$

$$\begin{aligned}&\Big|\Big\langle\nabla M(\bar{x}^{k-\tau}-x^\star),\,F_i(\bar{z}_t^k,Y_{i,t}^k)-F_i(\bar{x}^{k-\tau},Y_{i,t}^k)\Big\rangle\Big|\le\frac{L}{\ell_{cm}^2}\,\|\bar{x}^{k-\tau}-x^\star\|_c\,\|\bar{z}_t^k-\bar{x}^{k-\tau}\|_c,\\&\Big|\Big\langle\nabla M(\bar{x}^{k-\tau}-x^\star),\,\bar{F}_i(\bar{x}^{k-\tau})-\bar{F}_i(\bar{z}_t^k)\Big\rangle\Big|\le\frac{L}{\ell_{cm}^2}\,\|\bar{x}^{k-\tau}-x^\star\|_c\,\|\bar{z}_t^k-\bar{x}^{k-\tau}\|_c.\end{aligned}\tag{60}$$

Averaging over i and multiplying by α yields

$$\begin{aligned}&\frac{\alpha}{m}\sum_{i=1}^{m}\mathbb{E}\left[\begin{array}{l}\Big\langle\nabla M(\bar{x}^{k-\tau}-x^\star),\,F_i(\bar{z}_t^k,Y_{i,t}^k)-F_i(\bar{x}^{k-\tau},Y_{i,t}^k)\Big\rangle\\\quad+\Big\langle\nabla M(\bar{x}^{k-\tau}-x^\star),\,\bar{F}_i(\bar{x}^{k-\tau})-\bar{F}_i(\bar{z}_t^k)\Big\rangle\end{array}\right]\\&\le\frac{2\alpha L}{\ell_{cm}^2}\,\mathbb{E}[\|\bar{x}^{k-\tau}-x^\star\|_c\,\|\bar{z}_t^k-\bar{x}^{k-\tau}\|_c].\end{aligned}\tag{61}$$

Combining (56), (59), and (61) gives the following explicit (non-asymptotic) inequality for each inner step t:

$$\begin{aligned}\mathcal{T}_{k,t}^{\mathrm{mk}}\le&\frac{2\alpha\,L_{ms}}{\ell_{cs}^2}\,\mathbb{E}[\|\bar{z}_t^k-\bar{x}^{k-\tau}\|_c\,(LR_{k,t}+B)]+\frac{2\alpha\epsilon}{\ell_{cm}^2}\,\mathbb{E}[\|\bar{x}^{k-\tau}-x^\star\|_c\big(L\|\bar{x}^{k-\tau}-x^\star\|_c+B\big)]\\&+\frac{2\alpha L}{\ell_{cm}^2}\,\mathbb{E}[\|\bar{x}^{k-\tau}-x^\star\|_c\,\|\bar{z}_t^k-\bar{x}^{k-\tau}\|_c].\end{aligned}\tag{62}$$

The remaining task is to upper bound the delay drift $\|\bar{z}_t^k-\bar{x}^{k-\tau}\|_c$ using the deterministic recursion.

Lemma 4 (Bound for delay drift)**.** *Consider $\epsilon>0$ and the round-delay τ defined in* (51)*. For any $k\ge\tau, t\in\{0,1,\cdots,H-1\}$, define*

$$\Gamma_{k,\tau}:=\frac{B}{L}+\max_{k-\tau\le q\le k}\max_{0\le s\le H-1}\Big(R_{q,s}+S_{q,s}\Big).\tag{63}$$

Then,

$$\|\bar{z}_t^k-\bar{x}^{k-\tau}\|_c\le\alpha H(\tau+1)(2L+1)\Gamma_{k,\tau}.\tag{64}$$

Proof. Write $\bar{z}_t^k-\bar{x}^{k-\tau}=(\bar{z}_t^k-\bar{x}^k)+(\bar{x}^k-\bar{x}^{k-\tau})$. Then,

$$\|\bar{z}_t^k-\bar{x}^k\|_c=\left\|\sum_{s=0}^{t-1}(\bar{z}_{s+1}^k-\bar{z}_s^k)\right\|_c\le\sum_{s=0}^{t-1}\|\Delta_{k,s}\|_c\le t\cdot\alpha(2L+1)\Gamma_{k,\tau}\le\alpha H(2L+1)\Gamma_{k,\tau}.\tag{65}$$

We need to bound $\|\Delta_{k,s}\|_c$,

$$\begin{aligned}\frac{1}{\alpha}\|\Delta_{k,s}\|_c&=\left\|\frac{1}{m}\sum_{i=1}^{m}F_i(z_{i,s}^k,Y_{i,s}^k)-\bar{z}_s^k\right\|_c\\&=\left\|\frac{1}{m}\sum_{i=1}^{m}\big(F_i(z_{i,s}^k,Y_{i,s}^k)-F_i(x^\star,Y_{i,s}^k)\big)+\frac{1}{m}\sum_{i=1}^{m}\big(F_i(x^\star,Y_{i,s}^k)-x^\star\big)+(x^\star-\bar{z}_s^k)\right\|_c\\&\le\underbrace{\left\|\frac{1}{m}\sum_{i=1}^{m}\big(F_i(z_{i,s}^k,Y_{i,s}^k)-F_i(x^\star,Y_{i,s}^k)\big)\right\|_c}_{\text{(I)}}+\underbrace{\left\|\frac{1}{m}\sum_{i=1}^{m}\big(F_i(x^\star,Y_{i,s}^k)-x^\star\big)\right\|_c}_{\text{(II)}}+R_{k,s}.\end{aligned}$$

For (I),

$$\text{(I)} \le \frac{1}{m}\sum_{i=1}^{m}\|F_i(z_{i,s}^k, Y_{i,s}^k) - F_i(x^\star, Y_{i,s}^k)\|_c \le \frac{L}{m}\sum_{i=1}^{m}\|z_{i,s}^k - x^\star\|_c$$
$$\le \frac{L}{m}\sum_{i=1}^{m}\Big(\|z_{i,s}^k - \bar{z}_s^k\|_c + \|\bar{z}_s^k - x^\star\|_c\Big) = L(S_{k,s} + R_{k,s}).$$

For (II),

$$\text{(II)} \le \frac{1}{m}\sum_{i=1}^{m}\|F_i(x^\star, Y_{i,s}^k) - x^\star\|_c \le B.$$

Combining the above bounds yields

$$\|\Delta_{k,s}\|_c \le \alpha\Big((L+1)R_{k,s} + LS_{k,s} + B\Big) \le \alpha(2L+1)\Gamma_{k,\tau}. \tag{66}$$

For the within-round drift,

$$\|\bar{z}_t^k - \bar{x}^k\|_c = \left\|\sum_{s=0}^{t-1}(\bar{z}_{s+1}^k - \bar{z}_s^k)\right\|_c \le \sum_{s=0}^{t-1}\|\Delta_{k,s}\|_c \le \alpha H(2L+1)\Gamma_{k,\tau}. \tag{67}$$

Then, we consider bounding $\left\|\bar{x}^k - \bar{x}^{k-\tau}\right\|_c$,

$$\bar{x}^k - \bar{x}^{k-\tau} = \sum_{q=k-\tau}^{k-1}(\bar{x}^{q+1} - \bar{x}^q) = \sum_{q=k-\tau}^{k-1}(\bar{z}_H^q - \bar{z}_0^q), \tag{68}$$

and similarly,

$$\|\bar{z}_H^q - \bar{z}_0^q\|_c \le \sum_{s=0}^{H-1}\|\Delta_{q,s}\|_c \le H\cdot\alpha(2L+1)\Gamma_{k,\tau}, \quad \text{for each } q \in [k-\tau, k-1].$$

Therefore,

$$\|\bar{x}^k - \bar{x}^{k-\tau}\|_c \le \sum_{q=k-\tau}^{k-1}\|\bar{z}_H^q - \bar{z}_0^q\|_c \le \tau\cdot\alpha H(2L+1)\Gamma_{k,\tau}. \tag{69}$$

Combining the two parts gives the deterministic bound

$$\|\bar{z}_t^k - \bar{x}^{k-\tau}\|_c \le \alpha H(\tau+1)(2L+1)\Gamma_{k,\tau}. \tag{70}$$

□

Lemma 5 (Markov bias bound for one inner step)**.** *Let $\epsilon = \epsilon(\alpha)$ and the round-delay τ be defined in* (51)*. For any $k \ge \tau$ and $t \in \{0, 1, \ldots, H-1\}$, with $\Gamma_{k,\tau}$ defined in* (63)*, we have*

$$\mathcal{T}_{k,t}^{\mathrm{mk}} \le \alpha^2 H(\tau+1)\,(2L+1)\,L\left(\frac{2L_{ms}}{\ell_{cs}^2} + \frac{2}{\ell_{cm}^2}\right)\mathbb{E}[\Gamma_{k,\tau}^2] + \frac{2\alpha\epsilon L}{\ell_{cm}^2}\,\mathbb{E}[\Gamma_{k,\tau}^2]. \tag{71}$$

Proof. By Lemma 4,

$$\|\bar{z}_t^k - \bar{x}^{k-\tau}\|_c \le \alpha H(\tau+1)(2L+1)\Gamma_{k,\tau}.$$

To bound the remaining factors by $\Gamma_{k,\tau}$, we have

$$\|\bar{x}^{k-\tau} - x^\star\|_c = R_{k-\tau,0} \le \Gamma_{k,\tau}, \qquad R_{k,t} \le \Gamma_{k,\tau}. \tag{72}$$

$$LR_{k,t} + B \le L\Gamma_{k,\tau}. \tag{73}$$

Then

$$\frac{2\alpha\, L_{ms}}{\ell_{cs}^2}\,\mathbb{E}[\|\bar{z}_t^k - \bar{x}^{k-\tau}\|_c\,(LR_{k,t}+B)] \le \frac{2\alpha\, L_{ms}}{\ell_{cs}^2}\,\alpha H(\tau+1)(2L+1)\,L\,\mathbb{E}[\Gamma_{k,\tau}^2]. \tag{74}$$

Using (72)–(73), the second term is bounded by

$$\frac{2\alpha\epsilon}{\ell_{cm}^2}\,\mathbb{E}[\|\bar{x}^{k-\tau}-x^\star\|_c\big(L\|\bar{x}^{k-\tau}-x^\star\|_c+B\big)] \le \frac{2\alpha\epsilon L}{\ell_{cm}^2}\,\mathbb{E}[\Gamma_{k,\tau}^2]. \tag{75}$$

Finally, using (70) and (72), the third term is bounded by

$$\frac{2\alpha L}{\ell_{cm}^2}\,\mathbb{E}[\|\bar{x}^{k-\tau}-x^\star\|_c\,\|\bar{z}_t^k-\bar{x}^{k-\tau}\|_c] \le \frac{2\alpha L}{\ell_{cm}^2}\,\alpha H(\tau+1)(2L+1)\,\mathbb{E}[\Gamma_{k,\tau}^2]. \tag{76}$$

Summing these three bounds yields (71). □

A.4 Analysis of outer loop

We now combine the inner-step descent estimate, the Markovian bias estimate, and the consensus recursion. Throughout this subsection, let $k \ge \tau$, where $\epsilon(\alpha)$ and τ are defined in (51). We use $\Gamma_{k,\tau}$ as defined in (63).

First, we record the inner-loop bounds needed below. Assume

$$\alpha(L+1)H \le \frac{1}{4}, \qquad \alpha LH \le \frac{1}{4}. \tag{77}$$

Then Lemma 1 gives

$$\max_{0\le t\le H-1} S_{k,t} \le 2S_{k,0} + 4\alpha H\left(L\max_{0\le t\le H-1} R_{k,t} + B\right). \tag{78}$$

Moreover, using

$$\|\bar{z}_{t+1}^k - \bar{z}_t^k\|_c \le \alpha\left((L+1)R_{k,t} + LS_{k,t} + B\right),$$

we obtain

$$\max_{0\le t\le H-1} R_{k,t} \le R_{k,0} + \alpha H\left((L+1)\max_{0\le t\le H-1} R_{k,t} + L\max_{0\le t\le H-1} S_{k,t} + B\right).$$

Using $\alpha(L+1)H \le 1/4$ gives

$$\max_{0\le t\le H-1} R_{k,t} \le \frac{4}{3}\left(R_{k,0} + \alpha H\left(L\max_{0\le t\le H-1} S_{k,t} + B\right)\right). \tag{79}$$

Substituting (79) into (78), and then using $\alpha LH \le 1/4$, yields

$$\max_{0\le t\le H-1} S_{k,t} \le 3S_{k,0} + 8\alpha H\left(LR_{k,0} + B\right). \tag{80}$$

Consequently,

$$\begin{aligned}\sum_{t=0}^{H-1}\mathbb{E}[S_{k,t}^2] &\le H\mathbb{E}\left[\left(\max_{0\le t\le H-1} S_{k,t}\right)^2\right] \\ &\le 18H\,\mathbb{E}[S_{k,0}^2] + 512\alpha^2H^3L^2u_{cm}^2\,\mathbb{E}[\Phi_{k,0}] + 256\alpha^2H^3B^2.\end{aligned} \tag{81}$$

Here we used $R_{k,0}^2 \le 2u_{cm}^2\Phi_{k,0}$.

We next sum the one-step potential inequality in Lemma 3. Assume, in addition to (77), that

$$\frac{3L_{ms}(L+1)^2u_{cm}^2}{\ell_{cs}^2}\,\alpha \le \frac{\kappa}{2}, \qquad \frac{\alpha\kappa H}{2} \le 1. \tag{82}$$

Then

$$1-\alpha\kappa + \frac{3L_{ms}\alpha^2(L+1)^2u_{cm}^2}{\ell_{cs}^2} \le 1 - \frac{\alpha\kappa}{2},$$

and hence

$$\left(1-\alpha\kappa+\frac{3L_{ms}\alpha^2(L+1)^2u_{cm}^2}{\ell_{cs}^2}\right)^H\le 1-\frac{\alpha\kappa H}{4}.$$

Combining Lemma 3, Lemma 5, and (81) gives

$$\begin{aligned}\mathbb{E}[\Phi_{k+1,0}]\le&\left(1-\frac{\alpha\kappa H}{4}\right)\mathbb{E}[\Phi_{k,0}]\\&+\left(\frac{\alpha L^2u_{cm}^2}{2\kappa\ell_{cm}^4}+\frac{3L_{ms}\alpha^2L^2}{2\ell_{cs}^2}\right)\\&\quad\cdot\left(18H\,\mathbb{E}[S_{k,0}^2]+512\alpha^2H^3L^2u_{cm}^2\,\mathbb{E}[\Phi_{k,0}]+256\alpha^2H^3B^2\right)\\&+\left(\alpha^2H^2(\tau+1)(2L+1)L\left(\frac{2L_{ms}}{\ell_{cs}^2}+\frac{2}{\ell_{cm}^2}\right)+\frac{2\alpha\epsilon(\alpha)HL}{\ell_{cm}^2}\right)\mathbb{E}[\Gamma_{k,\tau}^2]\\&+\frac{3L_{ms}\alpha^2H}{2\ell_{cs}^2}B^2.\end{aligned}\tag{83}$$

If also

$$256\frac{\alpha^2H^2L^4u_{cm}^4}{\kappa\ell_{cm}^4}+768\frac{L_{ms}\alpha^3H^2L^4u_{cm}^2}{\ell_{cs}^2}\le\frac{\kappa}{8},\tag{84}$$

then the $\mathbb{E}[\Phi_{k,0}]$ contribution generated by the inner-loop disagreement is absorbed into the descent term, and (83) becomes

$$\begin{aligned}\mathbb{E}[\Phi_{k+1,0}]\le&\left(1-\frac{\alpha\kappa H}{8}\right)\mathbb{E}[\Phi_{k,0}]\\&+\alpha HL^2\left(\frac{9u_{cm}^2}{\kappa\ell_{cm}^4}+\frac{27\alpha L_{ms}}{\ell_{cs}^2}\right)\mathbb{E}[S_{k,0}^2]\\&+\left(\alpha^2H^2(\tau+1)(2L+1)L\left(\frac{2L_{ms}}{\ell_{cs}^2}+\frac{2}{\ell_{cm}^2}\right)+\frac{2\alpha\epsilon(\alpha)HL}{\ell_{cm}^2}\right)\mathbb{E}[\Gamma_{k,\tau}^2]\\&+\left(\frac{3L_{ms}\alpha^2H}{2\ell_{cs}^2}+\frac{128\alpha^3H^3L^2u_{cm}^2}{\kappa\ell_{cm}^4}+\frac{384L_{ms}\alpha^4H^3L^2}{\ell_{cs}^2}\right)B^2.\end{aligned}\tag{85}$$

The squared consensus error after one communication also follows from Lemma 2 and (80). Specifically,

$$\begin{aligned}\mathbb{E}[S_{k+1,0}^2]\le&\,18m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\mathbb{E}[S_{k,0}^2]+512m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\alpha^2H^2L^2u_{cm}^2\,\mathbb{E}[\Phi_{k,0}]\\&+256m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\alpha^2H^2B^2.\end{aligned}\tag{86}$$

It remains to control the delayed term $\Gamma_{k,\tau}^2$. Assume

$$\tau\alpha LH\le\frac{1}{4},\qquad\left(\left(1+\rho\frac{u_{2c}}{\ell_{2c}}\sqrt{m}\right)(1+3\alpha L)^H\right)^\tau\le 2.\tag{87}$$

Using the recursion

$$R_{q,s+1}+S_{q,s+1}\le(1+3\alpha L)(R_{q,s}+S_{q,s})+3\alpha B$$

inside each round, together with the communication bound for $S_{q+1,0}$, gives

$$\begin{aligned}\mathbb{E}[\Gamma_{k,\tau}^2]\le&\,576u_{cm}^2\,\mathbb{E}[\Phi_{k-\tau,0}]+288\,\mathbb{E}[S_{k-\tau,0}^2]\\&+\left(81e^{3/2}\left(1+\rho\frac{u_{2c}}{\ell_{2c}}\sqrt{m}\right)^2+18\right)\frac{B^2}{L^2}.\end{aligned}\tag{88}$$

Moreover, by Lemma 4,

$$\|\bar{x}^k-\bar{x}^{k-\tau}\|_c\le\tau\alpha H(2L+1)\Gamma_{k,\tau}.$$

Thus

$$\Phi_{k-\tau,0} \le 2\Phi_{k,0} + \frac{\tau^2\alpha^2H^2(2L+1)^2}{\ell_{cm}^2}\Gamma_{k,\tau}^2.$$

Taking expectations and using (88), if

$$576u_{cm}^2\frac{\tau^2\alpha^2H^2(2L+1)^2}{\ell_{cm}^2} \le \frac{1}{2}, \tag{89}$$

then

$$\begin{aligned}\mathbb{E}[\Phi_{k-\tau,0}] \le 4\mathbb{E}[\Phi_{k,0}] + 576\frac{\tau^2\alpha^2H^2(2L+1)^2}{\ell_{cm}^2}\mathbb{E}[S_{k-\tau,0}^2] \\ + \frac{2\tau^2\alpha^2H^2(2L+1)^2}{\ell_{cm}^2}\left(81e^{3/2}\left(1+\rho\frac{u_{2c}}{\ell_{2c}}\sqrt{m}\right)^2+18\right)\frac{B^2}{L^2}.\end{aligned} \tag{90}$$

The final combination uses the following Lyapunov weight:

$$c_S := 4\alpha HL^2\left(\frac{9u_{cm}^2}{\kappa\ell_{cm}^4} + \frac{27\alpha L_{ms}}{\ell_{cs}^2}\right). \tag{91}$$

For every $k \ge \tau$, define

$$V^k := \mathbb{E}[\Phi_{k,0}] + c_S\mathbb{E}\left[S_{k,0}^2\right] + c_S\mathbb{E}\left[S_{k-\tau,0}^2\right]. \tag{92}$$

We collect the step-size and network conditions used in the final combination. First, assume

$$\alpha(L+1)H \le \frac{1}{4}, \qquad \alpha LH \le \frac{1}{4}, \tag{93}$$

$$\frac{3L_{ms}(L+1)^2u_{cm}^2}{\ell_{cs}^2}\alpha \le \frac{\kappa}{2}, \qquad \frac{\alpha\kappa H}{2} \le 1, \tag{94}$$

$$256\frac{\alpha^2H^2L^4u_{cm}^4}{\kappa\ell_{cm}^4} + 768\frac{L_{ms}\alpha^3H^2L^4u_{cm}^2}{\ell_{cs}^2} \le \frac{\kappa}{8}, \tag{95}$$

$$\tau\alpha LH \le \frac{1}{4}, \qquad \left(\left(1+\rho\frac{u_{2c}}{\ell_{2c}}\sqrt{m}\right)(1+3\alpha L)^H\right)^\tau \le 2, \tag{96}$$

$$576u_{cm}^2\frac{\tau^2\alpha^2H^2(2L+1)^2}{\ell_{cm}^2} \le \frac{1}{2}, \tag{97}$$

$$18m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2 \le \frac{1}{2}. \tag{98}$$

Assume also the final absorption conditions

$$\begin{aligned}10240m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\alpha^3H^3L^4u_{cm}^2\left(\frac{9u_{cm}^2}{\kappa\ell_{cm}^4}+\frac{27\alpha L_{ms}}{\ell_{cs}^2}\right) \\ + 4608\alpha^2H^2(\tau+1)(2L+1)Lu_{cm}^2\left(\frac{L_{ms}}{\ell_{cs}^2}+\frac{1}{\ell_{cm}^2}\right) + \frac{4608\alpha^2HL^2L_{ms}u_{cm}^2\ell_{cs}^2}{\ell_{cm}^4} \le \frac{3\alpha\kappa H}{32},\end{aligned} \tag{99}$$

and

$$\begin{aligned}288\left(\alpha^2H^2(\tau+1)(2L+1)L\left(\frac{2L_{ms}}{\ell_{cs}^2}+\frac{2}{\ell_{cm}^2}\right)+\frac{2\alpha^2HL^2L_{ms}\ell_{cs}^2}{\ell_{cm}^4}\right) \\ + \frac{576\tau^2\alpha^2H^2(2L+1)^2}{\ell_{cm}^2}\left(512c_Sm\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\alpha^2H^2L^2u_{cm}^2\right. \\ \left.+576u_{cm}^2\left(\alpha^2H^2(\tau+1)(2L+1)L\left(\frac{2L_{ms}}{\ell_{cs}^2}+\frac{2}{\ell_{cm}^2}\right)+\frac{2\alpha^2HL^2L_{ms}\ell_{cs}^2}{\ell_{cm}^4}\right)\right) \le \frac{c_S}{4}.\end{aligned} \tag{100}$$

Denote

$$\begin{aligned}\Xi := \frac{48L_{ms}}{\kappa\ell_{cs}^2} + \alpha H^2 L^2 \Bigg(& \frac{4096u_{cm}^2}{\kappa^2\ell_{cm}^4} + \frac{12288\alpha L_{ms}}{\kappa\ell_{cs}^2} \\ & + \frac{65536m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2}{\kappa}\left(\frac{9u_{cm}^2}{\kappa\ell_{cm}^4} + \frac{27\alpha L_{ms}}{\ell_{cs}^2}\right)\Bigg) \\ & + \left(81e^{3/2}\left(1+\rho\frac{u_{2c}}{\ell_{2c}}\sqrt{m}\right)^2 + 18\right) \\ & \cdot \Bigg(\frac{64H\left(\tau+1\right)\left(2L+1\right)}{\kappa L}\left(\frac{L_{ms}}{\ell_{cs}^2} + \frac{1}{\ell_{cm}^2}\right) \\ & + \frac{64L_{ms}\ell_{cs}^2}{\kappa\ell_{cm}^4} + \alpha^2H^2\tau^2(2L+1)^2 \\ & \cdot \Bigg(\frac{131072m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\alpha H^2L^2u_{cm}^2}{\kappa\ell_{cm}^2} \\ & \cdot \left(\frac{9u_{cm}^2}{\kappa\ell_{cm}^4} + \frac{27\alpha L_{ms}}{\ell_{cs}^2}\right) \\ & + \frac{73728H\left(\tau+1\right)\left(2L+1\right)u_{cm}^2}{\kappa L\ell_{cm}^2}\left(\frac{L_{ms}}{\ell_{cs}^2} + \frac{1}{\ell_{cm}^2}\right) \\ & + \frac{73728L_{ms}u_{cm}^2\ell_{cs}^2}{\kappa\ell_{cm}^6}\Bigg)\Bigg). \end{aligned} \tag{101}$$

Proposition 1 (Round-level Lyapunov recursion for DLSA)**.** *Assume the hypotheses of Lemmas 3–5. If the stepsize and network parameter satisfy* (93)–(100)*, then, for every $k \geq \tau$,*

$$V^{k+1} \leq \left(1 - \frac{\alpha\kappa H}{32}\right)V^k + \frac{\alpha^2\kappa H}{32}B^2\Xi. \tag{102}$$

Proof. Add (85), c_S times (86), and c_S times the shifted version of (86). Then use (88) to replace $\mathbb{E}\left[\Gamma_{k,\tau}^2\right]$, and use (90) to replace $\mathbb{E}[\Phi_{k-\tau,0}]$. The coefficient of $\mathbb{E}[\Phi_{k,0}]$ is at most $1 - \alpha\kappa H/32$ by (99). By the definition of c_S, (98), and (94), the coefficient of $\mathbb{E}\left[S_{k,0}^2\right]$ is at most $(1 - \alpha\kappa H/32)c_S$. The coefficient of $\mathbb{E}\left[S_{k-\tau,0}^2\right]$ is at most $(1 - \alpha\kappa H/32)c_S$ by (98) and (100). In the remaining B^2 terms, use τ from (51) and collect the common leading factor $\alpha^2\kappa H/32$. These remaining terms are bounded by $(\alpha^2\kappa H/32)B^2\Xi$ according to (101). This proves (102). □

Corollary 3 (Finite-time convergence and steady-state error)**.** *Under the conditions of Proposition 1, with $R_k := \|\bar{x}^k - x^\star\|_c$ and $S_k := S_{k,0}$, for every $k \geq \tau + 1$,*

$$\mathbb{E}[R_k^2] = \widetilde{\mathcal{O}}\left(\left(1 - \frac{\alpha\kappa H}{32}\right)^{k-\tau}\left(\mathbb{E}[R_\tau^2] + \alpha H L^2 \mathbb{E}[S_\tau^2] + \alpha H L^2 \mathbb{E}[S_0^2]\right)\right) + \widetilde{\mathcal{O}}\left(\begin{array}{l}\alpha B^2\big(1 + \dfrac{H(2L+1)}{L} + \alpha H^2 L^2 \\ \quad + \alpha^2(1+H)^2(2L+1)^2\big(1 + \dfrac{H(2L+1)}{L} + \alpha H^2 L^2\big)\big)\end{array}\right), \tag{103}$$

$$\mathbb{E}[S_k^2] = \widetilde{\mathcal{O}}\left(\left(18m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\right)^{k-\tau}\mathbb{E}[S_\tau^2]\right) + \widetilde{\mathcal{O}}\left(\begin{array}{l}\rho^2\alpha^2 H^2 L^2\left(1 - \dfrac{\alpha\kappa H}{32}\right)^{k-\tau} \\ \quad \times \left(\mathbb{E}[R_\tau^2] + \alpha H L^2 \mathbb{E}[S_\tau^2] + \alpha H L^2 \mathbb{E}[S_0^2]\right)\end{array}\right) + \widetilde{\mathcal{O}}\left(\begin{array}{l}\rho^2\alpha^2 H^2 B^2\big(1 + \alpha L^2\big(1 + \dfrac{H(2L+1)}{L} + \alpha H^2 L^2 \\ \quad + \alpha^2(1+H)^2(2L+1)^2\big(1 + \dfrac{H(2L+1)}{L} + \alpha H^2 L^2\big)\big)\big)\end{array}\right). \tag{104}$$

Consequently,

$$\limsup_{k\to\infty} \mathbb{E}[R_k^2] \tag{105}$$

$$= \widetilde{\mathcal{O}}\left(\begin{array}{l}\alpha B^2\big(1 + \dfrac{H(2L+1)}{L} + \alpha H^2 L^2 \\ \quad + \alpha^2(1+H)^2(2L+1)^2\big(1 + \dfrac{H(2L+1)}{L} + \alpha H^2 L^2\big)\big)\end{array}\right), \tag{106}$$

$$\limsup_{k\to\infty} \mathbb{E}[S_k^2] \tag{107}$$

$$= \widetilde{\mathcal{O}}\left(\begin{array}{l}\rho^2\alpha^2 H^2 B^2\big(1 + \alpha L^2\big(1 + \dfrac{H(2L+1)}{L} + \alpha H^2 L^2 \\ \quad + \alpha^2(1+H)^2(2L+1)^2\big(1 + \dfrac{H(2L+1)}{L} + \alpha H^2 L^2\big)\big)\big)\end{array}\right). \tag{108}$$

Proof. Iterating (102) from τ to k gives

$$V^k \leq \left(1 - \frac{\alpha\kappa H}{32}\right)^{k-\tau} V^\tau + \alpha B^2 \Xi. \tag{109}$$

Since $R_k^2 \leq 2u_{cm}^2 \Phi_{k,0}$, this implies

$$\mathbb{E}\big[R_k^2\big] \leq 2u_{cm}^2\left(1 - \frac{\alpha\kappa H}{32}\right)^{k-\tau} V^\tau + 2u_{cm}^2 \alpha B^2 \Xi. \tag{110}$$

For the consensus error, iterate (86). For every $k \geq \tau + 1$,

$$\begin{aligned}\mathbb{E}\big[S_k^2\big] \leq{}& \left(18m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\right)^{k-\tau}\mathbb{E}\big[S_\tau^2\big] \\ &+ 512m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2 \alpha^2 H^2 L^2 u_{cm}^2 \sum_{j=\tau}^{k-1}\left(18m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\right)^{k-1-j}\mathbb{E}[\Phi_{j,0}] \\ &+ 256m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2 \alpha^2 H^2 B^2 \sum_{j=\tau}^{k-1}\left(18m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\right)^{k-1-j}.\end{aligned} \tag{111}$$

Using (109) to upper bound $\mathbb{E}[\Phi_{j,0}]$ and summing the two geometric series gives

$$\begin{aligned}\mathbb{E}\big[S_k^2\big] \le{}& \left(18m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\right)^{k-\tau}\mathbb{E}\big[S_\tau^2\big] \\ &+ \frac{512m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\alpha^2H^2L^2u_{cm}^2}{1-\frac{\alpha\kappa H}{32}-18m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2}\left(1-\frac{\alpha\kappa H}{32}\right)^{k-\tau-1}V^\tau \\ &+ \frac{512m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\alpha^3H^2L^2u_{cm}^2B^2\Xi+256m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2\alpha^2H^2B^2}{1-18m\rho^2\left(\frac{u_{2c}}{\ell_{2c}}\right)^2}. \end{aligned} \tag{112}$$

From (101), under the step-size conditions of Proposition 1, and after hiding norm-equivalence constants, κ, L_{ms}, Markov-chain constants, the effective network factor, and logarithmic factors in $\widetilde{\mathcal{O}}$, we have

$$\Xi = \widetilde{\mathcal{O}}\left(1+\frac{H(2L+1)}{L}+\alpha H^2L^2+\alpha^2(1+H)^2(2L+1)^2\left(1+\frac{H(2L+1)}{L}+\alpha H^2L^2\right)\right). \tag{113}$$

Moreover, by the definition of V^τ and c_S,

$$V^\tau = \widetilde{\mathcal{O}}\big(\mathbb{E}\big[R_\tau^2\big]+\alpha HL^2\mathbb{E}\big[S_\tau^2\big]+\alpha HL^2\mathbb{E}\big[S_0^2\big]\big). \tag{114}$$

Substituting (113) and (114) into (110) and (112) gives (103) and (104). Taking $\limsup_{k\to\infty}$ gives (106) and (108). □

Theorem 2 (Simplified sufficient conditions and DLSA convergence)**.** *Let τ be given by* (51)*, and let*

$$L_+ := \max\{L,1\}, \qquad \delta := \frac{u_{2c}}{\ell_{2c}}.$$

There exist positive constants c_α and c_ρ, depending only on the norm-equivalence, envelope, contraction, and Markov-chain constants, such that the conditions

$$\alpha \le c_\alpha \min\left(\frac{1}{HL_+},\frac{\kappa}{L_+^2},\frac{\kappa}{HL_+^2},\frac{\kappa^{1/3}}{H^{2/3}L_+^{4/3}},\frac{1}{\tau HL_+},\frac{\kappa}{H(\tau+1)L_+^2}\right) \tag{115}$$

and

$$\rho \le \frac{c_\rho}{\delta\sqrt{m}\,\tau} \tag{116}$$

imply (93)–(100)*. Consequently, for every $k \ge \tau+1$,*

$$\begin{aligned}\mathbb{E}\big[R_k^2\big] ={}& \widetilde{\mathcal{O}}\left(\left(1-\frac{\alpha\kappa H}{32}\right)^{k-\tau}\left(\mathbb{E}\big[R_\tau^2\big]+\alpha HL_+^2\mathbb{E}\big[S_\tau^2\big]+\alpha HL_+^2\mathbb{E}\big[S_0^2\big]\right)\right) \\ &+\widetilde{\mathcal{O}}\big(\alpha B^2\left(1+H+\alpha H^2L_+^2+\alpha^2(1+H)^2L_+^2\left(1+H+\alpha H^2L_+^2\right)\right)\big), \end{aligned} \tag{117}$$

$$\begin{aligned}\mathbb{E}\big[S_k^2\big] ={}& \widetilde{\mathcal{O}}\Big(\left(18m\rho^2\delta^2\right)^{k-\tau}\mathbb{E}\big[S_\tau^2\big]\Big) \\ &+\widetilde{\mathcal{O}}\left(\rho^2\alpha^2H^2L_+^2\left(1-\frac{\alpha\kappa H}{32}\right)^{k-\tau}\left(\mathbb{E}\big[R_\tau^2\big]+\alpha HL_+^2\mathbb{E}\big[S_\tau^2\big]+\alpha HL_+^2\mathbb{E}\big[S_0^2\big]\right)\right) \\ &+\widetilde{\mathcal{O}}\left(\begin{array}{l}\rho^2\alpha^2H^2B^2\big(1+\alpha L_+^2\big(1+H+\alpha H^2L_+^2 \\ \qquad+\alpha^2(1+H)^2L_+^2\big(1+H+\alpha H^2L_+^2\big)\big)\big)\end{array}\right). \end{aligned} \tag{118}$$

In particular,

$$\limsup_{k\to\infty}\mathbb{E}\big[R_k^2\big] = \widetilde{\mathcal{O}}\big(\alpha B^2\left(1+H+\alpha H^2L_+^2+\alpha^2(1+H)^2L_+^2\left(1+H+\alpha H^2L_+^2\right)\right)\big), \tag{119}$$

$$\limsup_{k\to\infty}\mathbb{E}\big[S_k^2\big] \tag{120}$$

$$= \widetilde{\mathcal{O}}\big(\rho^2\alpha^2H^2B^2\left(1+\alpha L_+^2\left(1+H+\alpha H^2L_+^2+\alpha^2(1+H)^2L_+^2\left(1+H+\alpha H^2L_+^2\right)\right)\right)\big). \tag{121}$$

Proof. We first verify that (115) and (116) imply the conditions of Proposition 1. The first term in (115) gives

$$\alpha(L+1)H = O(c_\alpha), \qquad \alpha LH = O(c_\alpha),$$

which implies (93) when c_α is small enough. The second term gives

$$\alpha(L+1)^2 = O(c_\alpha \kappa),$$

and hence implies the first part of (94). The bound $\alpha H L_+ \le O(c_\alpha)$ also gives $\alpha\kappa H \le O(c_\alpha)$ after decreasing c_α, because κ is fixed in the problem constants; hence the second part of (94) holds.

For the round-absorption condition (95), the third and fourth terms in (115) imply

$$\alpha^2 H^2 L^4 = O(c_\alpha^2 \kappa^2), \qquad \alpha^3 H^2 L^4 = O(c_\alpha^3 \kappa),$$

so the two terms on the left-hand side of (95) are at most constant multiples of $c_\alpha^2\kappa$ and $c_\alpha^3\kappa$. Thus (95) follows for sufficiently small c_α.

The fifth term in (115) gives

$$\tau\alpha LH \le \tau\alpha H L_+ \le O(c_\alpha), \qquad \tau\alpha H(2L+1) \le O(c_\alpha).$$

Together with (116), this gives

$$\tau \log\left(1 + \rho\delta\sqrt{m}\right) + \tau H \log(1 + 3\alpha L) \le O(c_\rho) + O(c_\alpha).$$

Taking c_α and c_ρ small enough makes the last display at most $\log 2$, which proves the second part of (96); the first part was already shown. The same bound $\tau\alpha H(2L+1) = O(c_\alpha)$ implies (97). Moreover, since $\tau \ge 1$, (116) gives

$$18 m \rho^2 \delta^2 \le O(c_\rho^2),$$

which implies (98) for sufficiently small c_ρ.

It remains to check the final absorption conditions. The left-hand side of (99) is bounded, up to fixed problem constants, by

$$\rho^2\alpha^3 H^3 L_+^4 + \alpha^2 H^2(\tau+1)L_+^2 + \alpha^2 H L_+^2.$$

Using (116), the first term is no larger than a constant multiple of $c_\rho^2 \alpha^3 H^3 L_+^4/\tau^2$. The third and sixth terms in (115) imply that all three displayed terms are $O(c_\alpha \alpha\kappa H)$. Thus (99) follows after decreasing c_α and c_ρ.

For (100), note that $c_S = \Theta(\alpha H L^2)$ up to fixed problem constants, and the left-hand side is a finite sum of monomials dominated by

$$\alpha^2 H^2(\tau+1)L_+^2, \qquad \alpha^2 H L_+^2, \qquad \tau^2\alpha^2 H^2 L_+^2 \left(\rho^2 c_S \alpha^2 H^2 L_+^2 + \alpha^2 H^2(\tau+1)L_+^2 + \alpha^2 H L_+^2\right).$$

The fifth and sixth terms in (115) imply $\tau\alpha H L_+ = O(c_\alpha)$ and $\alpha H(\tau+1)L_+^2 = O(c_\alpha\kappa)$. Together with (116), every monomial in the last display is at most a sufficiently small constant multiple of c_S. Therefore (100) holds.

Thus all conditions of Proposition 1 hold. Applying Corollary 3 gives the finite-time bounds. Finally, in the error terms of Corollary 3, substitute $2L+1 \le 3L_+$ and use $L_+ \ge 1$. The logarithmic factors coming from the ceiling expression in (101) are absorbed into $\widetilde{\mathcal{O}}$. This yields (117)–(121). □

B Specialization to Decentralized Q-learning

This section verifies that fixed-behavior tabular decentralized Q-learning is a special case of the decentralized stochastic approximation model in Section 3.

B.1 Q-learning as stochastic approximation

Let $d := |\mathcal{S}||\mathcal{A}|$. For each agent $i \in [m]$, the behavior policy π_i and transition kernel P_i induce a Markov chain over state-action pairs with transition matrix

$$K_i\big((s,a),(s',a')\big) = P_i(s' \mid s,a)\pi_i(a' \mid s'). \tag{122}$$

We assume that this chain is irreducible and aperiodic. Let $d_i(s,a)$ denote its stationary distribution. Since the state-action space is finite, there exist constants $c_{p,Q} > 0$ and $r_{p,Q} \in (0,1)$, uniform over agents, such that

$$d_{TV}\big(K_i^\ell((s,a),\cdot), d_i\big) \le c_{p,Q} r_{p,Q}^\ell, \qquad \forall i \in [m],\ \ell \ge 0. \tag{123}$$

The sample used by Q-learning is

$$Y_i^{(\ell)} = \big(S_i^{(\ell)}, A_i^{(\ell)}, R_i^{(\ell)}, S_i^{(\ell+1)}\big), \qquad R_i^{(\ell)} = r_i(S_i^{(\ell)}, A_i^{(\ell)}).$$

Since $Y_i^{(\ell)}$ is generated from the state-action chain through a fixed transition kernel, the same geometric mixing form holds for the sample chain, possibly after changing $c_{p,Q}$. Thus $r_{p,Q}$ is the Q-learning specialization of the Markov mixing rate r_p in (4).

For a sample $y = (s,a,r,s')$, define

$$F_i(Q,y) := Q + e_{s,a}\Big(r + \gamma_{\text{disc}} \max_{a' \in \mathcal{A}} Q(s',a') - Q(s,a)\Big), \tag{124}$$

where $e_{s,a} \in \mathbb{R}^d$ is the coordinate vector associated with (s,a). Then

$$Q + \alpha(F_i(Q,y) - Q) = Q + \alpha e_{s,a}\Big(r + \gamma_{\text{disc}} \max_{a' \in \mathcal{A}} Q(s',a') - Q(s,a)\Big), \tag{125}$$

which is exactly one sampled tabular Q-learning update at the visited state-action pair. Therefore Algorithm 1 with the operator (124) is decentralized Q-learning with H local sampled Bellman updates between two communication rounds.

For each agent, define the local Bellman optimality operator

$$(T_i Q)(s,a) := r_i(s,a) + \gamma_{\text{disc}} \sum_{s' \in \mathcal{S}} P_i(s' \mid s,a) \max_{a' \in \mathcal{A}} Q(s',a'). \tag{126}$$

Let

$$D_i := \operatorname{diag}\big(d_i(s,a) : (s,a) \in \mathcal{S} \times \mathcal{A}\big), \qquad \bar{d}(s,a) := \frac{1}{m}\sum_{i=1}^m d_i(s,a),$$

and

$$\bar{D} := \operatorname{diag}\big(\bar{d}(s,a) : (s,a) \in \mathcal{S} \times \mathcal{A}\big).$$

Assume

$$d_{\min} := \min_{(s,a) \in \mathcal{S}\times\mathcal{A}} \bar{d}(s,a) > 0. \tag{127}$$

Taking expectation of (124) under the stationary sample distribution of agent i gives

$$\bar{F}_i(Q) = \mathbb{E}_{Y \sim \mu_i^{\pi_i}}[F_i(Q,Y)] = Q + D_i(T_i Q - Q). \tag{128}$$

Averaging over agents,

$$\bar{F}_Q(Q) := \frac{1}{m}\sum_{i=1}^m \bar{F}_i(Q) = Q + \frac{1}{m}\sum_{i=1}^m D_i(T_i Q - Q). \tag{129}$$

Define

$$(\widetilde{T}Q)(s,a) := \frac{\sum_{i=1}^m d_i(s,a)(T_i Q)(s,a)}{\sum_{i=1}^m d_i(s,a)}. \tag{130}$$

Since $d_{\min} > 0$, the denominator is positive for every (s,a). Then (129) becomes

$$\bar{F}_Q(Q) = Q + \bar{D}(\widetilde{T}Q - Q). \tag{131}$$

The operator $\widetilde{T}$ is a discounted Bellman optimality operator with reward and transition kernel

$$\widetilde{r}(s,a) = \frac{\sum_{i=1}^m d_i(s,a) r_i(s,a)}{\sum_{i=1}^m d_i(s,a)}, \qquad \widetilde{P}(\cdot \mid s,a) = \frac{\sum_{i=1}^m d_i(s,a) P_i(\cdot \mid s,a)}{\sum_{i=1}^m d_i(s,a)}.$$

Thus $\widetilde{T}$ is a γ_{disc}-contraction in $\|\cdot\|_\infty$ and has a unique fixed point, denoted by $Q^\star$. Moreover, since $\bar{D}$ is positive diagonal,

$$\bar{F}_Q(Q) = Q \quad \Longleftrightarrow \quad \bar{D}(\widetilde{T}Q - Q) = 0 \quad \Longleftrightarrow \quad \widetilde{T}Q = Q. \tag{132}$$

Hence the fixed point of the averaged stochastic approximation operator $\bar{F}_Q$ is exactly $Q^\star$.

B.2 Verification of Assumption 1

We verify Assumption 1 with

$$\|\cdot\|_c = \|\cdot\|_\infty.$$

For any $y = (s, a, r, s')$ and $Q, Q' \in \mathbb{R}^d$, the sampled operator satisfies

$$\|F_i(Q, y) - F_i(Q', y)\|_\infty \le \|Q - Q'\|_\infty. \tag{133}$$

Also,

$$\begin{aligned} &F_i(Q, y) - Q - F_i(Q', y) + Q' \\ &= e_{s,a}\left(\gamma_{\text{disc}}\left[\max_{a'} Q(s', a') - \max_{a'} Q'(s', a')\right] - (Q(s, a) - Q'(s, a))\right), \end{aligned} \tag{134}$$

and therefore

$$\|F_i(Q, y) - Q - F_i(Q', y) + Q'\|_\infty \le (1 + \gamma_{\text{disc}})\|Q - Q'\|_\infty. \tag{135}$$

Thus Assumption 1(i) holds with

$$L_Q := 1 + \gamma_{\text{disc}}. \tag{136}$$

Next, using (131),

$$\bar{F}_Q(Q) - \bar{F}_Q(Q') = (I - \bar{D})(Q - Q') + \bar{D}(\widetilde{T}Q - \widetilde{T}Q').$$

Since $\widetilde{T}$ is a γ_{disc}-contraction in $\|\cdot\|_\infty$, for every (s, a),

$$\begin{aligned} |(\bar{F}_Q(Q) - \bar{F}_Q(Q'))(s, a)| &\le (1 - \bar{d}(s, a))\|Q - Q'\|_\infty + \bar{d}(s, a)\gamma_{\text{disc}}\|Q - Q'\|_\infty \\ &= \left(1 - (1 - \gamma_{\text{disc}})\bar{d}(s, a)\right)\|Q - Q'\|_\infty \\ &\le \left(1 - (1 - \gamma_{\text{disc}})d_{\min}\right)\|Q - Q'\|_\infty. \end{aligned} \tag{137}$$

Therefore Assumption 1(ii) holds with

$$\beta_Q := 1 - (1 - \gamma_{\text{disc}})d_{\min}. \tag{138}$$

Finally, since $Q^\star$ is the fixed point of a discounted Bellman optimality operator with rewards bounded by $R_{\max}$,

$$\|Q^\star\|_\infty \le \frac{R_{\max}}{1 - \gamma_{\text{disc}}}. \tag{139}$$

For any sample $y = (s, a, r, s')$,

$$\|F_i(Q^\star, y)\|_\infty \le \max\left\{\|Q^\star\|_\infty,\, R_{\max} + \gamma_{\text{disc}}\|Q^\star\|_\infty\right\} \le \frac{R_{\max}}{1 - \gamma_{\text{disc}}}.$$

Thus Assumption 1(iii) holds with

$$B_Q := \frac{2R_{\max}}{1 - \gamma_{\text{disc}}}. \tag{140}$$

B.3 Moreau-envelope constants for the ℓ_∞ contraction

The Moreau-envelope analysis requires the constants relating $\|\cdot\|_\infty$ and the Moreau-envelope norm $\|\cdot\|_M$. Following the generalized Moreau-envelope construction, choose

$$p_Q := \max\{2, 2\log d\}, \qquad \|\cdot\|_s := \|\cdot\|_{p_Q}. \tag{141}$$

Then

$$d^{-1/p_Q}\|x\|_{p_Q} \le \|x\|_\infty \le \|x\|_{p_Q}, \qquad \forall x \in \mathbb{R}^d. \tag{142}$$

Thus

$$\ell_{cs,Q} = d^{-1/p_Q}, \qquad u_{cs,Q} = 1. \tag{143}$$

Take

$$\theta_Q := \left(\frac{1 + \beta_Q}{2\beta_Q}\right)^2 - 1. \tag{144}$$

By the generalized Moreau-envelope norm equivalence,

$$\ell_{cm,Q} = \left(1 + \theta_Q \ell_{cs,Q}^2\right)^{1/2} = \left(1 + \theta_Q d^{-2/p_Q}\right)^{1/2}, \tag{145}$$

and

$$u_{cm,Q} = \left(1 + \theta_Q u_{cs,Q}^2\right)^{1/2} = (1 + \theta_Q)^{1/2} = \frac{1 + \beta_Q}{2\beta_Q}. \tag{146}$$

The smoothness constant of the envelope may be taken as

$$L_{ms,Q} := \frac{p_Q - 1}{\theta_Q}. \tag{147}$$

Therefore the contraction gap in the Moreau-envelope Lyapunov recursion is

$$\kappa_Q := 1 - \beta_Q \frac{u_{cm,Q}}{\ell_{cm,Q}} = 1 - \beta_Q \left(\frac{1 + \theta_Q}{1 + \theta_Q d^{-2/p_Q}}\right)^{1/2}. \tag{148}$$

Equivalently,

$$\kappa_Q = 1 - \frac{1 + \beta_Q}{2}\left(1 + d^{-2/p_Q}\left[\left(\frac{1 + \beta_Q}{2\beta_Q}\right)^2 - 1\right]\right)^{-1/2}. \tag{149}$$

When $d \geq 3$, $p_Q = 2\log d$, and hence $d^{-2/p_Q} = e^{-1}$. In this common case,

$$\kappa_Q = 1 - \frac{1 + \beta_Q}{2}\left(1 + \frac{1}{e}\left[\left(\frac{1 + \beta_Q}{2\beta_Q}\right)^2 - 1\right]\right)^{-1/2}. \tag{150}$$

B.4 Applying the DLSA theorem

We now apply Theorem 1. In the Q-learning specialization, the generic quantities R_k and S_k in Theorem 1 are read under the substitutions

$$x_i^k = Q_i^k, \qquad x^\star = Q^\star, \qquad \|\cdot\|_c = \|\cdot\|_\infty.$$

Thus the theorem controls the squared estimation error of the averaged Q-table and the squared consensus error among local Q-tables.

Define the Q-learning Markov delay

$$\tau_Q := \max\left\{1, \left\lceil \frac{1}{H(1 - r_{p,Q})} \log \frac{c_{p,Q}\ell_{cm,Q}^2}{L_Q L_{ms,Q}\ell_{cs,Q}^2 \alpha} \right\rceil\right\}. \tag{151}$$

Then

$$\tau_Q = \widetilde{\mathcal{O}}\left(1 + \frac{1}{H(1 - r_{p,Q})}\right). \tag{152}$$

Let ρ_{eff} denote the effective network mixing factor in one communication round. For plain DLSA, $\rho_{\text{eff}} = \rho$; for DLSA-FastMix, $\rho_{\text{eff}} = \rho_{\text{fm}}$. If

$$\alpha \leq \frac{c_{\alpha,Q}}{H(\tau_Q + 1)(1 + L_Q)^2}, \qquad \rho_{\text{eff}} \leq \frac{c_{\rho,Q}}{\sqrt{m}(\tau_Q + 1)} \tag{153}$$

for sufficiently small problem-dependent constants $c_{\alpha,Q}, c_{\rho,Q} > 0$, then Theorem 1 gives, for all $k \geq \tau_Q$,

$$\begin{aligned}\mathbb{E}\left[R_k^2\right] &\leq C_{R,Q}\left(1 - \frac{\alpha\kappa_Q H}{32}\right)^{k-\tau_Q}\left(\mathbb{E}\left[R_{\tau_Q}^2\right] + \mathbb{E}\left[S_{\tau_Q}^2\right] + \mathbb{E}\left[S_0^2\right]\right) \\ &\quad + \widetilde{\mathcal{O}}\left(\alpha B_Q^2\left(1 + H + \alpha H^2(1 + L_Q)^2 + \alpha^2 H^3(1 + L_Q)^2\right)\right), \end{aligned} \tag{154}$$

$$\begin{aligned}\mathbb{E}\left[S_k^2\right] &\leq C_{S,Q}\left(1 - \frac{\alpha\kappa_Q H}{32}\right)^{k-\tau_Q}\left(\mathbb{E}\left[R_{\tau_Q}^2\right] + \mathbb{E}\left[S_{\tau_Q}^2\right] + \mathbb{E}\left[S_0^2\right]\right) \\ &\quad + \widetilde{\mathcal{O}}\left(\rho_{\text{eff}}^2\alpha^2 H^2 B_Q^2\left[1 + \alpha(1 + L_Q)^2\left(1 + H + \alpha H^2(1 + L_Q)^2 + \alpha^2 H^3(1 + L_Q)^2\right)\right]\right). \end{aligned} \tag{155}$$

Here $C_{R,Q}, C_{S,Q} > 0$ are problem-dependent constants, and the hidden constants in $\widetilde{\mathcal{O}}(\cdot)$ depend only on fixed norm-equivalence, envelope, Markov-chain, and contraction constants.

Corollary 1 now applies with the substitutions

$$L = L_Q, \qquad B = B_Q, \qquad \tau = \tau_Q, \qquad c_0 = \frac{\kappa_Q}{32}.$$

For a target accuracy $\varepsilon \in (0, 1)$, choose an admissible constant stepsize satisfying

$$\alpha = \min\left\{ \frac{c_{\alpha,Q}}{H(\tau_Q + 1)(1 + L_Q)^2}, \frac{\varepsilon}{B_Q^2(1 + H)}, \frac{\sqrt{\varepsilon}}{B_Q H} \right\}. \tag{156}$$

Then after

$$K = \widetilde{\mathcal{O}}\left(\tau_Q + \frac{1}{\alpha \kappa_Q H} \log \frac{1}{\varepsilon} \right) \tag{157}$$

communication rounds,

$$\mathbb{E}\big[R_K^2\big] \le \widetilde{\mathcal{O}}(\varepsilon), \qquad \mathbb{E}\big[S_K^2\big] \le \widetilde{\mathcal{O}}(\varepsilon). \tag{158}$$

The corresponding local sample complexity per agent is

$$KH = \widetilde{\mathcal{O}}\left(\tau_Q H + \frac{1}{\alpha \kappa_Q} \log \frac{1}{\varepsilon} \right). \tag{159}$$

Equivalently, substituting (156) gives the number of communication rounds

$$\widetilde{\mathcal{O}}\left(\tau_Q + \max\left\{ \frac{(\tau_Q + 1)(1 + L_Q)^2}{\kappa_Q}, \frac{B_Q^2(1 + H)}{\kappa_Q H \varepsilon}, \frac{B_Q}{\kappa_Q \sqrt{\varepsilon}} \right\} \log \frac{1}{\varepsilon} \right), \tag{160}$$

and the local sample complexity

$$\widetilde{\mathcal{O}}\left(\tau_Q H + \max\left\{ \frac{H(\tau_Q + 1)(1 + L_Q)^2}{\kappa_Q}, \frac{B_Q^2(1 + H)}{\kappa_Q \varepsilon}, \frac{B_Q H}{\kappa_Q \sqrt{\varepsilon}} \right\} \log \frac{1}{\varepsilon} \right). \tag{161}$$

In particular, when H, L_Q, B_Q are fixed and $\alpha = \Theta(\varepsilon)$ is admissible, the leading-order numbers of communication rounds and samples per agent reduce to

$$\widetilde{\mathcal{O}}\left(\tau_Q + \frac{1}{\kappa_Q H \varepsilon} \log \frac{1}{\varepsilon} \right) \quad \text{and} \quad \widetilde{\mathcal{O}}\left(\tau_Q H + \frac{1}{\kappa_Q \varepsilon} \log \frac{1}{\varepsilon} \right), \tag{162}$$

respectively. Multiplying the round count by 1 for plain DLSA or by n_ρ for DLSA-FastMix gives the communication complexity in gossip steps.

B.5 Local Q-functions and the common target

The common target $Q^\star$ solves the aggregate Bellman equation, while each agent evaluates rewards and transitions in its own MDP. Let $\pi_{\mathrm{agg}}(s) \in \arg\max_a Q^\star(s, a)$, using any fixed rule to break ties. Let $Q_i^\star$ be agent i's optimal Q-function and $Q_i^{\pi_{\mathrm{agg}}}$ its Q-function when following π_{agg}. To quantify their differences from $Q^\star$, define

$$\eta_i = \|(T_i - \widetilde{T})Q^\star\|_\infty, \qquad \Delta_{r,i} = \|r_i - \widetilde{r}\|_\infty, \qquad \Delta_{P,i} = \max_{s,a} \|P_i(\cdot \mid s, a) - \widetilde{P}(\cdot \mid s, a)\|_1,$$

where $\widetilde{r}$ and $\widetilde{P}$ are the aggregate rewards and transitions defined above. The quantity η_i is the discrepancy between the local and aggregate Bellman updates evaluated at the common target.

Proposition 2 (Local and aggregate Q-functions)**.** *The tabular setting above gives*

$$\max\{\|Q_i^\star - Q^\star\|_\infty, \|Q_i^{\pi_{\mathrm{agg}}} - Q^\star\|_\infty\} \le \frac{\eta_i}{1 - \gamma_{\mathrm{disc}}}, \tag{163}$$

$$\eta_i \le \Delta_{r,i} + \frac{\gamma_{\mathrm{disc}} R_{\max}}{1 - \gamma_{\mathrm{disc}}} \Delta_{P,i}. \tag{164}$$

The local policy loss also obeys

$$\|V_i^\star - V_i^{\pi_{\mathrm{agg}}}\|_\infty \le \frac{2\eta_i}{(1 - \gamma_{\mathrm{disc}})^2}.$$

Proof. Since $\widetilde{T}Q^\star = Q^\star$, contraction of T_i yields

$$\|Q_i^\star - Q^\star\|_\infty \le \gamma_{\mathrm{disc}}\|Q_i^\star - Q^\star\|_\infty + \eta_i.$$

For policy evaluation, let $T_i^{\pi_{\mathrm{agg}}}$ denote the Bellman operator with the next-state action fixed to π_{agg}. Greediness gives $T_i^{\pi_{\mathrm{agg}}}Q^\star = T_iQ^\star$. The same contraction calculation therefore applies to $Q_i^{\pi_{\mathrm{agg}}}$, proving (163). Expanding $(T_i - \widetilde{T})Q^\star$ into reward and transition differences and using $\|Q^\star\|_\infty \le R_{\max}/(1-\gamma_{\mathrm{disc}})$ gives (164). Finally, with $\delta_i = \|Q_i^\star - Q^\star\|_\infty$, greediness implies $V_i^\star(s) - Q_i^\star(s, \pi_{\mathrm{agg}}(s)) \le 2\delta_i$. The Bellman equation for $V_i^{\pi_{\mathrm{agg}}}$ then gives a policy loss at most $2\delta_i/(1-\gamma_{\mathrm{disc}})$. □

For example, consider one state, two actions a, b, two agents, and $\gamma_{\mathrm{disc}} = 1/2$. Both agents remain in the same state, with reward vectors $r_1 = (1, 0)$ and $r_2 = (0, 3)$. Uniform coverage gives $Q^\star = (2, 3)$ and the aggregate greedy policy chooses b. Locally, $Q_1^\star = (2, 1)$, $Q_1^{\pi_{\mathrm{agg}}} = (1, 0)$, and $Q_2^\star = Q_2^{\pi_{\mathrm{agg}}} = (3, 6)$. Thus the aggregate policy is not optimal for agent 1. Both agents have $\eta_i = 3/2$ and $\|Q_i^{\pi_{\mathrm{agg}}} - Q^\star\|_\infty = 3$, attaining the bound $\eta_i/(1-\gamma_{\mathrm{disc}})$.

For a finite iterate, the triangle inequality separates the errors as

$$\|Q_i^k - Q_i^\star\|_\infty \le \|Q_i^k - \bar{Q}^k\|_\infty + \|\bar{Q}^k - Q^\star\|_\infty + \frac{\eta_i}{1-\gamma_{\mathrm{disc}}}.$$

The first two terms are consensus and estimation errors relative to the common target. The last term reflects heterogeneity between the aggregate and local MDPs.

C Additional Experiments

We first describe the Q-learning instances and then give the linear-SA and Markovian-SGD studies.

C.1 Q-learning instances and agent heterogeneity

The agents share finite state and action spaces but have different rewards and transition kernels. For each agent i and state-action pair (s, a), independently generate

$$P_i(\cdot \mid s, a) \sim \mathrm{Dirichlet}(0.3\mathbf{1}_{|\mathcal{S}|}), \qquad r_i(s, a) \sim \mathrm{Unif}[-2, 2].$$

Problem seed 11 is used for each instance, drawing all transition rows before all rewards. The generated MDPs are fixed across the compared methods and values of H. At each local update, (s, a) is drawn independently and uniformly, followed by $s' \sim P_i(\cdot \mid s, a)$. Thus the Q-learning experiments use an i.i.d. generative sampler, which is a special case of the Markovian sampling model, rather than a behavior-policy trajectory. In particular, $d_i(s, a) = 1/(|\mathcal{S}||\mathcal{A}|)$ and the reference $Q^\star$ solves the Bellman equation of the uniformly averaged MDP with $\bar{P} = m^{-1}\sum_i P_i$ and $\bar{r} = m^{-1}\sum_i r_i$.

The first instance has $(m, |\mathcal{S}|, |\mathcal{A}|, \gamma_{\mathrm{disc}}) = (10, 8, 4, 0.95)$, and the FastMix instance has $(8, 4, 2, 0.9)$. Hence $d_{\min} = 1/32$ and $1/8$, respectively. The path graph uses Metropolis weights $w_{ij} = 1/(1 + \max\{\deg(i), \deg(j)\})$ on each edge and $w_{ii} = 1 - \sum_{j \ne i} w_{ij}$, where $\deg(i)$ is the degree of node i.

To summarize the generated heterogeneity, define

$$\Delta_r = \left[\frac{1}{m|\mathcal{S}||\mathcal{A}|}\sum_{i,s,a}(r_i(s, a) - \bar{r}(s, a))^2\right]^{1/2},$$
$$\Delta_P = \frac{1}{m|\mathcal{S}||\mathcal{A}|}\sum_{i,s,a}\frac{1}{2}\|P_i(\cdot \mid s, a) - \bar{P}(\cdot \mid s, a)\|_1.$$

The two instances give $(\Delta_r, \Delta_P) = (1.092, 0.496)$ and $(1.079, 0.473)$ after rounding. Table 3 reports the mean and maximum of the individual quantities defined in Appendix B.5. These were computed from the disclosed generator by solving the aggregate, local-optimal, and fixed-policy Bellman equations to residuals below 10^{-12}. The transition statistic Δ_P is a mean total-variation distance, while each $\Delta_{P,i}$ is a maximum full ℓ_1 distance.

Table 3: Heterogeneity of the two MDP instances. Means and maxima are over agents.

Quantity	$m = 10,\ \lvert\mathcal{S}\rvert = 8,\ \lvert\mathcal{A}\rvert = 4$ Mean	Maximum	$m = 8,\ \lvert\mathcal{S}\rvert = 4,\ \lvert\mathcal{A}\rvert = 2$ Mean	Maximum
$\Delta_{r,i}$	2.1402	2.7259	1.7473	2.3867
$\Delta_{P,i}$	1.4704	1.6471	1.3778	1.6878
η_i	2.1115	2.7275	1.7763	2.4347
$\lVert Q_i^\star - Q^\star\rVert_\infty$	16.1232	25.8120	9.2130	17.1877
$\lVert Q_i^{\pi_{\text{agg}}} - Q^\star\rVert_\infty$	8.5753	18.2480	8.8002	16.6642

C.2 Linear SA

We first consider a decentralized linear-SA / quadratic-optimization problem with network-average objective

$$f(x) = \frac{1}{m}\sum_{i=1}^{m} f_i(x),$$

where each local loss $f_i(x) = \mathbb{E}[\ell_i(x; Y_{i,t}^k)]$ is accessed only through a stream of temporally correlated Markovian samples $\{Y_{i,t}^k\}$. Each agent forms a stochastic gradient $g_i(x; Y_{i,t}^k)$ and aims to converge to the unique minimizer $x^\star$. The stepsize is constant within each run

but is tuned separately for each H. The finite-horizon search returns the largest tested feasible stepsize α_H at its final bracket resolution, as detailed below. We report

$$\text{err}_k = \frac{1}{m}\sum_{i=1}^{m} \lVert x_i^k - x^\star\rVert_2$$

at each communication round. Here each communication round consists of H local stochastic gradient steps $z_{i,t+1}^k = z_{i,t}^k - \alpha_H g_i(z_{i,t}^k; Y_{i,t}^k)$ for $t = 0, \ldots, H-1$, followed by one consensus step $X^{k+1} = (W \otimes I_d) Z_H^k$ using a symmetric doubly-stochastic mixing matrix W on a connected peer-to-peer graph. We measure communication cost by the number of times W is applied, which cleanly separates the effect of the local-update parameter H in error-versus-communication plots. Here we fix the number of agents to $m = 10$ and the dimension to $d = 50$. We consider $H \in \{1, 3, 6, 9, 12\}$ and run each method for $K = 1500$ communication rounds.

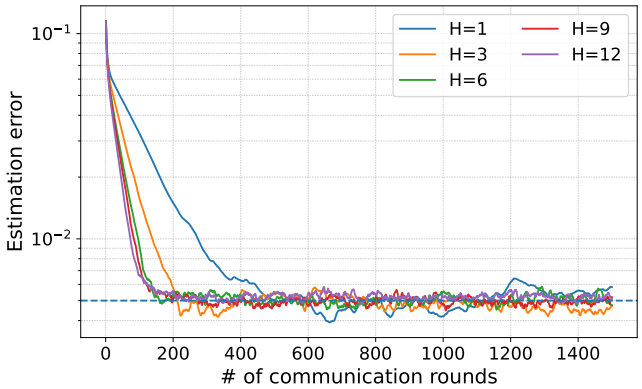


Figure 4: Estimation error versus communication rounds for different numbers of local updates H in the linear-SA experiment. The blue dashed line denotes the target precision 5×10^{-3}.

Figure 4 illustrates that more local updates can reduce the communication needed to reach the target error over the tested range of H.

C.2.1 Stepsize selection

For each H, the stepsize search uses $K = 1500$ communication rounds. A candidate is feasible if its mean estimation error over the final 200 rounds is at most 5×10^{-3}. Starting from $\alpha_{\max} = 0.02$, we halve the stepsize until a feasible candidate is found. This candidate and the preceding larger infeasible candidate define the search interval. Arithmetic bisection tests their midpoint and replaces the feasible lower endpoint or the infeasible upper endpoint accordingly. We stop after at most 20 bisections or when the interval width is below 10^{-5}, returning the feasible lower endpoint. If the initial value 0.02 is feasible, we retain it as the tested cap. Candidates for a fixed H share the same initialization and Markov-noise seed. This finite-horizon search is a tuning heuristic, since stochastic feasibility need not be globally monotone in the stepsize.

C.2.2 Stepsize sensitivity

To complement Figure 4, we study sensitivity on a fully specified supplementary linear-SA instance. The experiment fixes $m = 10$, $d = 50$, and a path graph with Metropolis weights. For indices $i = 0, \dots, 9$ and $j = 0, \dots, 49$, set

$$a_{ij} = 0.8\,\mathbf{1}\{i + j \text{ is even}\}, \qquad A_i = \operatorname{diag}(a_{i0}, \dots, a_{i49}).$$

Problem seed 11 generates a vector $x^\star$ of Euclidean norm 0.1 and unit vectors v_h, v_i by normalizing independent Gaussian draws. With $q_i = (-1)^i$, define

$$h_i = 0.1[(I - W)q]_i v_h, \qquad b_i = A_i x^\star + h_i, \qquad F_i(x, s) = x - A_i x + b_i - 0.03 s\, v_i.$$

For each agent, $s \in \{-1, +1\}$ follows an independent stationary Markov chain with probability 0.9 of retaining its sign. Since $\sum_i h_i = 0$ and $m^{-1} \sum_i A_i = 0.4I$, the average mean operator is a 0.6-contraction with fixed point $x^\star$. Each local mean operator leaves some coordinates unchanged and is not a strict contraction. All iterates start at zero.

For $H \in \{1, 3, 6, 9, 12\}$, we select α_H by the search above using tuning seed 1001 and then test $\alpha = c\alpha_H$ with $c \in \{0.5, 0.75, 1, 1.25, 1.5\}$. Each run has 1500 communication rounds. Evaluation uses five independent stationary noise streams with seeds 2001–2005, sharing the same realized stream across candidates within each seed. The supplementary instance gives

H	1	3	6	9	12
α_H	0.019873	0.018848	0.010625	0.007871	0.006221

We report the error averaged over the final 200 rounds and the first communication round T at which the mean local Euclidean error is at most 0.005. A run that does not reach the target is right-censored at 1500, so the communication panel summarizes $\min\{T, 1500\}$. Points and error bars are means and one sample standard deviation across the five evaluation streams. The instance, initialization, budget, and seeds are fixed across the stepsize comparisons.

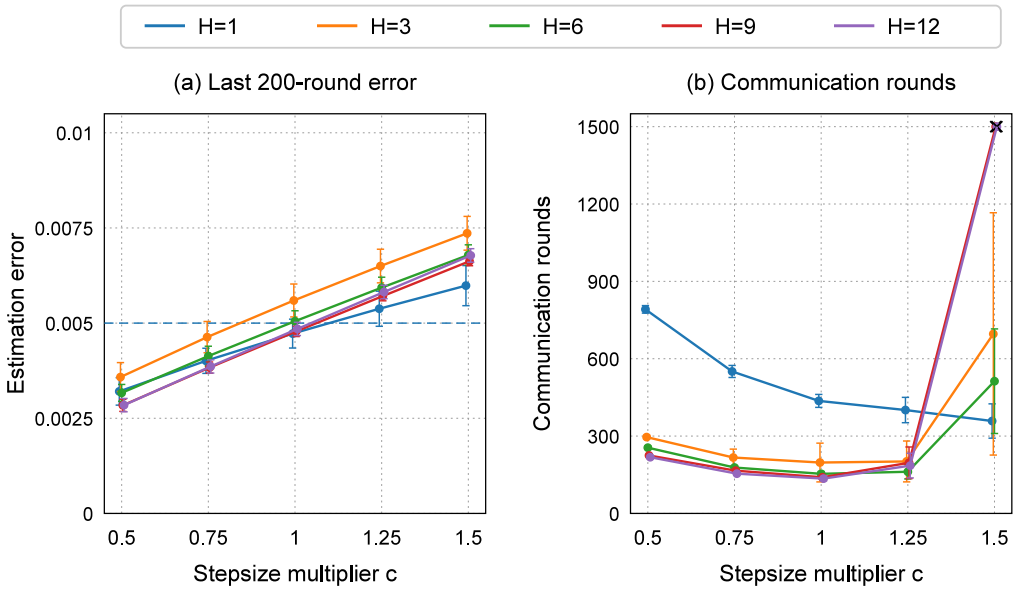


Figure 5: Stepsize sensitivity in the supplementary linear-SA instance. The panels report mean error over the final 200 rounds and restricted first-hit communication time. Error bars show one sample standard deviation across five streams. Black crosses mark settings with a censored run. The blue dashed line is the target error, and small horizontal offsets separate overlapping error bars.

Of the 125 evaluation runs, 115 reach the target within the budget. At $c = 1$, mean first-hit times decrease from 436.4 rounds for $H = 1$ to 135.0 for $H = 12$. The $H = 3$ mean terminal error is nevertheless about 0.00559, so a feasible stepsize on the tuning stream need not attain the same final accuracy on independent streams. At $c = 1.5$, all five runs for each of $H = 9$ and $H = 12$ fail to reach the target. These results support reporting final error together with communication time. The sampled operators satisfy the finite-state and bounded-noise assumptions, but the empirical settings are not claimed to satisfy every conservative stepsize and network threshold of the theorem. The supplementary code includes all tuning trials, full curves, seeds, and censoring indicators.

The supplementary study ran on an AMD Ryzen 5 6600H CPU with 32 GB of host memory and no GPU, using Python 3.12.14, NumPy 2.3.5, and ReportLab 4.4.9. The measured time for all 60 tuning candidates, 125 evaluation runs, and CSV output was about 13.1 seconds, excluding plotting and document generation.

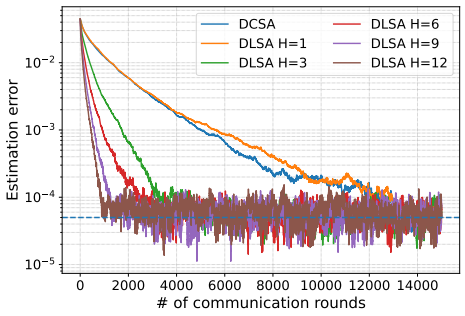
C.3 Decentralized local SGD with Markovian samples

To better evaluate the effectiveness of DLSA, we construct a decentralized system-identification problem with a synthetic Markovian data model. We initialize $m = 50$ agents, each maintaining a $d = 20$-dimensional Markov state $\phi_{i,t} \in \mathbb{R}^d$. For each agent i, at time t we draw $u_{i,t} \sim \mathcal{N}(0, 1)$, set the first coordinate $\phi_{i,t}(1) = u_{i,t}$, and update the remaining coordinates via the subdiagonal recursion

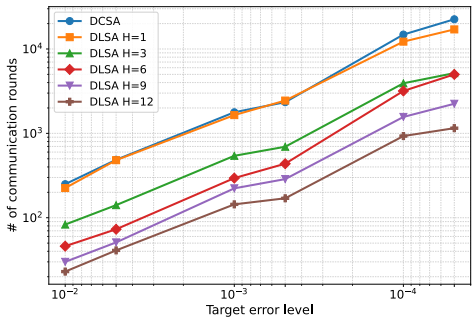
$$\phi_{i,t}(r) = a_{i,r}\, \phi_{i,t-1}(r-1), \qquad r = 2, \ldots, d,$$

where $a_{i,r} \sim \mathrm{Unif}[0.8, 0.99]$. A global unknown parameter $x^\star$ is sampled uniformly from the ℓ_2 unit ball, and each agent forms squared-loss stochastic gradients from the resulting Markovian observations. Communication is defined on a line graph using a lazy Metropolis mixing matrix W.

Given this streaming Markov data, we compare DLSA with different H values against the $H = 1$ baseline DCSA [41] to estimate $x^\star$. For each method and each H, we automatically select a stable constant stepsize via a feasibility search so that all methods reach a common target steady-state error. The implementation tracks performance by forming the network average $\bar{x}^k = \frac{1}{m}\sum_{i=1}^m x_i^k$ after each communication round and recording the estimation error $\|\bar{x}^k - x^\star\|_2$.



(a) Estimation error versus communication rounds for different numbers of local updates H for Markovian SGD. The blue dashed line denotes the target precision, which is set to 5×10^{-5}.



(b) Communication Rounds Required to Reach a Target Estimation Error.

Figure 6: Communication–Accuracy Tradeoff under Different Numbers of Local Updates for Markovian SGD

The results in Figure 6(a) show that when $H = 1$, DCSA performs similarly to DLSA, but as H increases, the advantage of doing more local updates becomes apparent. DLSA achieves the same accuracy with fewer communication rounds, thereby significantly reducing the communication cost. Besides, we also report the number of communication rounds needed to reach each target error level. The curves compare the baseline DCSA with DLSA using different numbers of local updates $H \in \{1, 3, 6, 9, 12\}$. For each method and each H, we use the same stepsize-selection procedure and report the smallest round index k such that the target error is achieved.

As shown in Figure 6(b), local SA requires fewer consensus operations than DCSA to reach the same target accuracy, and the gap widens as the target becomes more stringent. While DCSA effectively pays a communication cost at every noisy update, local SA amortizes communication by performing $H > 1$ local SGD steps before mixing. Over the tested range of H, this leads to a better communication–accuracy tradeoff. More broadly, however, the theory does not predict that larger H is always better: overly large local-update budgets or poorly tuned stepsizes can increase disagreement and worsen the eventual error floor. The Gaussian feature model has unbounded sampled gradients and lies outside the finite-state and uniform sampled-operator Lipschitz assumptions of the theorem.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The main claims made in the abstract and introduction accurately reflect the paper's contributions and scope.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The paper's limitations have been characterized in the assumption statements and in the discussions of future work in the conclusion.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: The paper's assumption statements are written out clearly.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [No]

   Justification: The supplementary linear-SA study in Appendix C.2.2 is fully specified. The complete implementation and raw traces of the original comparison figures are not included, so those figures cannot all be reproduced from the supplied material.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [No]

Justification: The accompanying source archive includes code, generated instances, and raw results for the supplementary studies. A public repository and the complete original experimental code are not provided.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [No]

   Justification: Appendix C specifies the supplementary instances, search rules, and evaluation seeds. Exact run configurations are not available for every original comparison figure.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [No]

   Justification: Figure 5 reports means and sample standard deviations over five independent noise streams. The original comparison figures do not report uncertainty across independent runs.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [No]

   Justification: Appendix C.2.2 reports the CPU, memory, software versions, and measured runtime of the supplementary study. Equivalent records are not available for all original experiments.

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer:[Yes]

   Justification: This paper is primarily theoretical.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A]

    Justification: This paper is primarily theoretical.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.

- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: This paper is primarily theoretical.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A]

    Justification: This paper is primarily theoretical.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.

- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: The supplementary source archive documents the generated instances, parameter settings, seeds, full curves, and censoring indicators described in Appendix C.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: This paper is primarily theoretical.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: This paper is primarily theoretical.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.

- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: LLMs assisted manuscript editing and supplementary code preparation. They are not a component of the DLSA algorithm or the sampling model.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.