Finite-Time Convergence Rates of Decentralized Stochastic Approximation With Applications in Multi-Agent and Multi-Task Learning

Sihan Zeng, Graduate Student Member, IEEE, Thinh T. Doan, Member, IEEE, and Justin Romberg, Fellow, IEEE

Abstract—In this article, we study a decentralized variant of stochastic approximation (SA), a data-driven approach for finding the root of an operator under noisy measurements. A network of agents, each with its own operator and data observations, cooperatively find the fixed point of the aggregate operator over a decentralized communication graph. Our main contribution is to provide a finite-time analysis of this decentralized SA method when the data observed at each agent are sampled from a Markov process; this lack of independence makes the iterates biased and (potentially) unbounded. Under fairly standard assumptions, we show that the convergence rate of the proposed method is essentially the same as if the samples were independent, differing only by a log factor that accounts for the mixing time of the Markov processes. The key idea in our analysis is to introduce a novel Lyapunov–Razumikhin function, motivated by the one used in analyzing the stability of delayed ordinary differential equations. We also discuss applications of the proposed method on a number of interesting learning problems in multiagent systems.

Index Terms—Distributed optimization, reinforcement learning, stochastic approximation (SA).

I. INTRODUCTION

STochastic approximation (SA), introduced in [1], is a sample-based approach for finding the root (or fixed point) of some unknown operator under corrupted measurements. This method has broad applications in many areas including stochastic optimization [2] and reinforcement learning [3], as existing algorithms in these areas can be viewed as different variants of SA. In this article, our interest is to study a decentralized stochastic approximation (DSCA) framework, where there are a group of $N$ agents interacting through a decentralized communication network. The goal of the agents is to collaboratively find the root of a global operator composed of local operators distributed among the agents.

Mathematically, each agent $i$ has access to a local operator $F_i: X_i \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, where $X_i$ is a statistical sample space with distribution $\mu_i$. The goal of the agents is to find $\bar{\theta}$ satisfying

$$\bar{F}(\bar{\theta}) \triangleq \sum_{i=1}^{N} F_i(\bar{\theta}) = 0$$

where $\bar{F}_i$ is defined as

$$\bar{F}_i(\theta) \triangleq \mathbb{E}_{\mu_i}[F_i(X_i, \theta)] = \sum_{X_i \in X_i} \mu_i(X_i) F_i(X_i, \theta).$$

We assume that at least one solution $\bar{\theta}$ to (1) exists, and the structure of $\bar{F}$ that we discuss later implies that $\bar{\theta}$ is unique. The agents can interact with each other through an undirected and connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where agents $i, j \in \mathcal{V}$ communicate if and only if $(i, j) \in \mathcal{E}$. We denote by $\mathcal{N}(i)$ the neighboring set of agent $i$, i.e., $\mathcal{N}(i) \triangleq \{ j \mid (i, j) \in \mathcal{E}\}$.

We are motivated by the wide applications of decentralized algorithms under the rapid development of wireless sensor networks, mobile robots, and the Internet of Things. In these areas, since the data are already collected in a decentralized manner, decentralized approaches for solving the problem are preferred over their centralized counterparts. For example, when the data dimension is very large (e.g., for image or video sensors) transmitting the raw data to a fusion center for processing is time consuming and requires massive communication. In addition, mobile sensors usually have limited communication capacity and are only able to communicate with other agents nearby, for example, in the application of using unmanned robots with limited energy to collaboratively explore an unknown terrain or boundary [4].

Our proposed DCSA for solving (1) is formally stated in Algorithm 1, where each agent $i$ maintains $\theta_i^k$, a local estimate of the optimal $\bar{\theta}$. In every iteration $k$, the iterate $\theta_i^k$ is updated according to (2), a combination between a consensus step and a local SA step. In the consensus step, the agent $i$ collects its neighbors parameters $\{\theta_j^k \mid j \in \mathcal{N}_i\}$ and performs a weighted average using the weights $W(i, j)$ to achieve an agreement

---

1 We use “agent” to represent for a sensor, a robot, or a smart device.
between the agents’ iterates. On the other hand, to push the consensus point toward θ∗ agent i moves in the direction of its local operator $F_i(X_i^k, \theta_i^k)$ scaled by some step size $\epsilon_k$.

Our focus is to study the finite-time performance of DCSA when the data $\{X_i^k\}$ observed at each agent i are sampled from Markov processes. This Markov sampling results in the data being dependent and biased. The existing theoretical guarantees for the finite-time performance of DCSA are studied under the common assumption that the local data at each agent are independent and identically distributed (i.i.d.). Such an assumption may not hold in many applications where the data are dependent and evolve over time, for example, when they are sampled from dynamical systems like in stochastic control or reinforcement learning.

### A. Main Contribution

The main contribution of this work is to establish the finite-time convergence rates of DCSA (Algorithm 1) when the data at each agent are generated by Markov processes. We do so under general conditions on the local operators $F_i$ (Lipschitz continuity) and the global operator (1-point strongly monotone). In particular, we provide an explicit formula that characterizes the convergence rates of Algorithm 1 under constant and time-varying step sizes. Our results, explained in details in Section II, differ from the i.i.d. case by only a log factor, which naturally arises due to the geometric mixing time of the Markov chain. In addition, our rate matches the one in centralized SA with Markovian data up to a factor that depends on the structure of the communication graph $G$. This article significantly extends our preliminary work in [5] as discussed in details in Section I-B.

Key to our analysis is a novel Lyapunov function tailored to decentralized algorithms with Markovian samples. Since the data are dependent, the current iterates depend on the history of the algorithm. Thus, the analysis of Algorithm 1 must carefully account for data across different delays. Inspired by the popular use of Lyapunov–Razumikhin functions in studying the stability of delayed ordinary differential equations, we design a Lyapunov function of this type to derive the convergence rates for DCSA. We believe that this Lyapunov approach is new in the literature of decentralized algorithms, and is of independent interest as it may be applied to many problems in this setting.

Finally, we discuss how our theoretical results can be applied to characterize the performance of two important special cases: decentralized Markov chain gradient descent and decentralized Q-learning, and provide some illustrative numerical examples.

### B. Related Work

In this section, we provide a brief summary of the existing literature in decentralized stochastic gradient descent (DSGD) and decentralized reinforcement learning methods, both of which can be viewed as special cases of DCSA. This will help to distinguish this article with the existing work in these areas.

In [6], [7], [8], [9], and the references therein, DSGD is studied for optimizing convex and nonconvex problems under i.i.d. data (realizations of the gradient). In [10] and [11], Markovian data are considered, similar to the setting in this article. The existing theoretical results of DSGD are established based on various assumptions on the boundedness of the gradients/updates. For example, the authors in [10] assume that the norm of $F_i(X_i, \theta)$ is uniformly bounded, while a bounded variance of $\sum_{i=1}^{N}(F_i(X_i, \theta) - F_i(\theta))$ is assumed in [11]. Similarly, the bounded variance condition, a bound on the norm of the error between the stochastic gradient and the true gradient, is considered in [6]. The authors in [8] relax this condition to only require that the expected norm of the stochastic gradient is bounded affinely by the norm of the true gradient. However, [12], Prop. 1 shows that none of these conditions in [6], [7], [8], [9], [10], and [11] holds on a simple objective function. Relaxing these strong conditions is nontrivial, especially in the context of Markovian data and decentralized framework. One of the main contributions of this article is to loosen these assumptions considerably: we only assume that $F_i$ (which is the gradient map in DSGD) is Lipschitz continuous, which is observed by the example in [12], Prop. 1.

In the context of reinforcement learning, SA is a common tool to analyze the convergence of temporal-difference (TD) learning [13], Q-learning [14], and other value-based methods [3]. In the distributed setting, decentralized TD learning using linear function approximation has been studied in [15] and [16], which can be viewed as a special case of DCSA with $F_i$ being linear operators. Moreover, the authors in [16] consider the setting where the linear operators at the agents are the same. The techniques in [15] and [16] cannot be directly applied to analyze DCSA, since we consider nonlinear SA and the local operators are different across the agents. Moreover, DCSA captures the decentralized variant of the popular Q-learning under linear function approximation, whose convergence property is unknown in the existing literature. With an assumption on the state visitation, we show that the problem falls under our analytical framework and its convergence rate can be derived as a consequence of our main theorem. In addition, we note that the decentralized policy gradient methods for multiagent and/or multitask reinforcement learning in [17], [18], [19], [20], [21], and [22] also perform updates similar to (2) in spirit that combine local gradient steps with network-level information aggregation.

We also want to point out some related work in the parameter/server settings (or federated learning/local SGD), where the agents communicate with a centralized coordinator, a problem that has attracted a good deal of interest lately because of its broad applications in data centers and cloud computing. This centralized communication setting can be a special case of Algorithm 1 with a “star graph,” and the recent advances in its theoretical understanding can be found in [23], [24], [25], and the references therein. Extending these theoretical results in the centralized communication setting to the decentralized setting here, however, is not straightforward.
Finally, we note that a preliminary version of this work is available in [5], where we considered DCSA algorithm under a constant step size. In this work, we study both constant and time-varying step sizes and provide a complete analysis for the finite-time convergence rates of DCSA, which is missing in [5]. We also discuss our results when the underlying communication graphs are time varying, which is not covered in [5].

II. FinitE-Time Analysis

In this section, we present the main results of this article, where we explicitly characterize the convergence rate of Algorithm 1 in solving (1). We provide an upper bound on the error $\mathbb{E}[\|\theta_k^i - \theta^*\|^2]$ for any $i \in V$ when Algorithm 1 uses constant and time-varying step size sequences $\{\epsilon_k\}$. When $\epsilon_k = \epsilon$ for some constant $\epsilon$, this error obeys

$$
\mathbb{E}[\|\theta_k^i - \theta^*\|^2] \leq O\left(a^k + \epsilon \log\left(\frac{1}{\epsilon}\right)\right)
$$

for a constant $\alpha < 1$ explicitly shown later in the section. Under-time-varying step sizes $\epsilon_k = O(1/k)$

$$
\mathbb{E}[\|\theta_k^i - \theta^*\|^2] \leq O\left(\frac{\log^2(k)}{k}\right).
$$

We begin by introducing the technical assumptions and useful notation used in our analysis in the following sub-sections.

A. Technical Assumptions and Notation

Given a constant $\epsilon > 0$, we denote by $\tau(i, \epsilon)$ the mixing time associated with the Markov chain $\{X^k_i\}$, i.e., $\tau(i, \epsilon)$ is the smallest integer such that $\forall x_i \in X_i$

$$
d_TV(P(X^k_i = \cdot | X^0_i = x_i, \mu_i) \leq \epsilon \quad \forall k \geq \tau(i, \epsilon)
$$

where $d_TV(\cdot, \cdot)$ is the total variance distance and $P(X^k_i = \cdot | X^0_i = x_i)$ is the distribution of $X^k_i$ with the initial conditions $x_i$. The mixing time represents the time it takes for the distribution of $X^k_i$ to get close to the stationary distribution $\mu_i$. We consider the following technical assumption on the properties of the Markov chains $\{X^k_i\}$.

Assumption 1: The space $X_i$ are compact, and the Markov chains $\{X^k_i\}_k$ are ergodic for all $i = 1, 2, \ldots, N$.

Assumption 1 implies that the Markov chain $\{X^k_i\}_k$ has a geometric mixing time. In other words, there exist $m_i > 0$ and $\rho_i \in (0, 1)$ such that $\forall x_i \in X_i$

$$
d_TV(P(X^k_i = \cdot | X^0_i = x_i, \mu_i) \leq m_i \rho_i^k \quad \forall k \geq 0.
$$

In the special case where $X_i$ is finite, the ergodicity of the Markov chain is guaranteed when it is irreducible and aperiodic, and the mixing time depends on the second largest eigenvalue of the transition probability matrix of $\{X^k_i\}_k$ [26].

Equation (3) implies that there exists a constant $\beta > 0$ such that

$$
\tau(i, \epsilon) \leq \beta \log(1/\epsilon) \quad \forall i.
$$

Given any $\epsilon > 0$ and denoting $\rho \triangleq \max_{i=1, 2, \ldots, N} \rho_i$, we define

$$
\tau(\epsilon) \triangleq \max\left\{\frac{\rho}{1 - \rho}, \max_{i=1, 2, \ldots, N} \tau(i, \epsilon)\right\}.
$$

For convenience, we denote $\tau_k = \tau(\epsilon_k)$. We require $\tau(\epsilon) \geq \frac{1}{\rho}$ as this condition later helps us to show $\tau_k \leq \tau_{k-1} + 1$ for $k > \tau_k$, i.e., $\tau_k$ does not increase too fast. In addition, as long as the step size $\epsilon_k$ is a polynomial function of $k$, it is obvious that $\tau_k$ grows at most logarithmically in $k$, and we can find a constant $c_\epsilon \in (0, 1)$ such that

$$
\tau_k + 1 \leq (1 - c_\epsilon)(k + 1) \quad \forall k > \tau_k.
$$

Assumption 1 holds in various applications, e.g., in incremental optimization [27] and in reinforcement learning modeled by Markov decision processes with a finite number of states and actions. In addition, Assumption 1 is used in the existing literature to study the finite-time performance of SA and its distributed variants under Markov randomness; see [13] and [14] and the references therein.

Next, we consider two assumptions on the operators $F_i$.

Assumption 2: The mappings $F_i(X_i, \cdot)$ are Lipschitz continuous; there exists a constant $L > 0$ such that for all $x_i \in X_i$

$$
\|F_i(x_i, \theta) - F_i(x_i, \tilde{\theta})\| \leq L\|\theta - \tilde{\theta}\| \quad \forall \theta, \tilde{\theta} \in \mathbb{R}^d.
$$

Assumption 3: The mapping $-F$ is 1-point strongly monotone w.r.t $\theta^*$, i.e., there exists $\alpha > 0$ such that

$$
(F(\theta) - F(\tilde{\theta}), \theta - \tilde{\theta}) \leq -\alpha\|\theta - \tilde{\theta}\|^2 \quad \forall \theta, \tilde{\theta} \in \mathbb{R}^d.
$$

As a result of Assumption 2, the mapping $F_i(\cdot)$ is also Lipschitz continuous with constant $L$.

$$
\|F_i(\theta) - F_i(\tilde{\theta})\| \leq L\|\theta - \tilde{\theta}\| \quad \forall \theta, \tilde{\theta} \in \mathbb{R}^d.
$$

Since $X_i$ are compact we denote by

$$
B \triangleq \max\{L, \max_{i \in V, x_i \in X_i} \|F_i(x_i, 0)\|\} < \infty.
$$

Under Assumption 3, $F$ is 1-point strongly monotone, which implies that the solution $\theta^*$ of (1) is unique. This assumption obviously holds in the case where the operator $F$ is globally strongly monotone, i.e., $\theta^*$ in (7) can be replaced by any $\theta^* \in \mathbb{R}^d$.

The global strong monotonicity is often assumed in the existing literature in studying the finite-time convergence rate of SA; see, for example, [14] and the references therein. In linear SA where the operator $F(\theta) = A\theta$ for some matrix $A \in \mathbb{R}^{d \times d}$, $A$ is usually assumed to be negative definite, which implies the strong monotonicity of $F$ [13], [16]. In the context of SGD, where $F$ is a gradient mapping, this assumption is equivalent to the the restricted Secant inequality [28], which is a generalized variant of strong convexity and is often used to guarantee the linear convergence of SGD for nonconvex optimization.

Finally, we consider the following assumption of the weight matrix of the network $W$, which encodes the communication structures shared between agents.

Assumption 4: The matrix of consensus weights $W = [W(i, j)] \in \mathbb{R}^{N \times N}$ in (2) is doubly stochastic, i.e., $\sum_i W_{ij} = 1$. In addition, $W(i, i) > 0$ and $W(i, j) > 0$ if and only if $(i, j) \in E$. Otherwise $W(i, j) = 0$.

This assumption is a standard condition in the existing literature of decentralized consensus-based algorithms to guarantee that the agents reach an agreement [29]. When the underlying network is connected, it implies that the largest singular value of $W$ is 1, while other singular values are strictly less than 1. The rates of Algorithm 1 depend on the spectral gap $1 - \sigma_2$ representing the connectivity of the network, where $\sigma_2 \in [0, 1)$ denotes the second largest singular value of $W$. In the rest of this article, we assume that Assumptions 1–4 always hold.

Notation: We use $\|\cdot\|$ and $\|\cdot\|_\infty$ to denote the matrix Frobenius norm and the spectral norm, respectively. We denote by $I$ and $\mathbf{1}$ the all-one vector in $\mathbb{R}^N$ and the identity matrix in $\mathbb{R}^{N \times N}$. 

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Given a sequence of local vectors \( \theta_1, \ldots, \theta_N \in \mathbb{R}^d \), we denote by \( \Theta \) a matrix in \( \mathbb{R}^{N \times d} \), whose \( i \)th row is \( \theta_i^T \). We then denote their average by \( \bar{\theta} \)

\[
\bar{\theta} \triangleq \frac{1}{N} \sum_{i=1}^{N} \theta_i = \frac{1}{N} \Theta^T 1 \in \mathbb{R}^d.
\]

### B. Main Results

We now present the main result of this article, where we provide an explicit formula to characterize the finite-time convergence rates of Algorithm 1 under both constant and time-varying step sizes. Recalling the definitions of \( L, \alpha, c_\tau, \) and \( B \) from Assumptions 2 and 3 and (6) and (8), we introduce the following constants for convenience.

\[
C_1 = \left( 60B^2 + \frac{45}{2} + 90BL + 6B^2 \right) \left( \|\theta^*\|^2 + 1 \right)
\]

\[
C_2 = \left( \frac{21B}{2} + 5 + \frac{8L^2}{\alpha} + 10L \right)
\]

\[
C_{c,1} = \max \left\{ 6B, \frac{45B + 132B^2 + 192BL}{\alpha} \right\}
\]

\[
C_{c,2} = \max \left\{ 16B, \frac{768B^2}{c_\tau \alpha}, \frac{\alpha}{4} + \frac{128B^2}{c_\tau (1 - \sigma_2^2)} + 2C_2, \frac{32B^2}{C_2} \right\}.
\]

Later in Section II-C, we will introduce the notion of optimality error and consensus error. Loosely speaking, \( C_1 \) describes the coupling between the two errors and \( C_1 \) represents the effect of the Markov randomness on the optimality error.

We consider the following two step sizes.

1) **Constant step size**: Let \( \epsilon_k = \epsilon \) (\( \tau_k = \tau(\epsilon) \)) satisfy

\[
\epsilon \tau(\epsilon) \leq \min \left\{ \frac{1}{N C_{c,1}}, \frac{1 - \sigma_2^2}{N C_{c,2}} \right\}.
\]

2) **Diminishing step size**: Let \( \epsilon_k \) be chosen as

\[
\epsilon_k = \frac{\epsilon}{k + 1}, \quad \text{where} \quad \epsilon \geq \frac{8}{\alpha}
\]

and \( \epsilon_k \tau_k \leq \min \left\{ \frac{1}{N C_{c,1}}, \frac{1 - \sigma_2^2}{N C_{c,2}} \right\} \quad \forall k \geq \tau_k. \quad \text{(11)}

By (4) we have \( \lim_{\epsilon \to 0} \epsilon \tau(\epsilon) = 0 \), so one can choose an \( \epsilon \) small enough that (10) is satisfied. A similar argument holds for (11).

**Theorem 1**: Let \( \{\theta_i^k\} \forall i \in \mathcal{V} \), be generated by Algorithm 1.

1) Under the constant step size in (10), we have \( \forall k \geq \tau(\epsilon) \)

\[
\mathbb{E} \left[ \|\theta_i^k - \theta^*\|^2 \right] \leq \left( 1 - \frac{8 \epsilon}{\alpha} \right)^{k - \tau(\epsilon)} \frac{4S^0 + 7\|\bar{\theta}^0\|^2 + C_1}{(1 - \sigma_2^2)}
\]

\[
+ \frac{208N \beta C_1}{\alpha(1 - \sigma_2^2)} \epsilon \log(\frac{1}{\epsilon}).
\]

2) Under the diminishing step size in (11), we have \( \forall k \geq \tau_k \)

\[
\mathbb{E} \left[ \|\theta_i^k - \theta^*\|^2 \right] \leq \frac{4(\beta \log(k + 1) - \log(\epsilon)) + 1}{(1 - \sigma_2^2)(k + 1)} \left( 4S^0 + 7\|\bar{\theta}^0\|^2 + C_1 \right)
\]

\[
+ \frac{208N \beta C_1 \log(k) \log(k - \log(\epsilon))}{\epsilon^2(1 - \sigma_2^2)(k + 1)}.
\]  

where \( c_\tau \in (0, 1) \) is defined in (6), and \( S^0 \) is a constant defined later in (14) that depends on the initial condition.

With its proof presented in Section IV, Theorem 1 states that under a constant step size \( \epsilon \) every local iterate generated by Algorithm 1 converges linearly to a ball around the optimal solution with radius on the order of \( \epsilon \log(\frac{1}{\epsilon}) \). Under a properly decaying step size, every local iterate converges exactly to the optimal solution with rate \( O(\log^2(k)/k) \).

If \( F_i \) is the gradient of some function, our result resembles the best known convergence rate of decentralized SGD for solving strongly convex objective functions under i.i.d. samples, differing only by a log factor (for example, see [30], Th. 2). In addition, this rate matches the ones in [25] where the authors consider a centralized communication topology.

Our bound depends inversely on \( 1 - \sigma_2^2 \), which reflects the impact of the graph \( \mathcal{G} \) and the matrix \( W \) on the convergence. The bound also scales proportional to \( N \), which shows the increasing difficulty of the problem as the number of agents goes up.

### C. Main Technical Challenges

One may attempt to apply the existing analysis in both decentralized optimization and reinforcement learning literature [16], [29] to derive the convergence rates presented in Theorem 1. However, due to the general settings we consider, in particular, the interaction between the decentralized framework and nonlinear SA under Markovian data and (potentially) unbounded operators, such an extension is nontrivial. In the following, we present the main technical challenges in our analysis and explain in detail the reason why the existing technique fails. We then discuss the novel idea in our analysis, where we introduce a variant of Lyapunov–Razumikhin functions [31]. Our complete analysis is given in Section IV.

Our goal is to study the convergence of every iterate \( \theta_i^k \) to the fixed point \( \theta^* \). Due to the consensus update, one can separate the error \( \theta_i^k - \theta^* \) as

\[
\theta_i^k - \theta^* = \bar{\theta}_i^k - \bar{\theta}^k + \bar{\theta}^k - \theta^k
\]

“consensus error” “optimality error” and study the convergence of these two errors to zero separately, as often done in the decentralized optimization literature [29]. However, there are two main challenges that make this approach inapplicable to our setting. For notation simplicity, we define the squared norm of the optimality error in iteration \( k \) as

\[
R^k \triangleq \|\bar{\theta}^k - \theta^*\|^2
\]

and the variance of consensus error in iteration \( k \) as

\[
S^k \triangleq \sum_{i=1}^{N} \|\theta_i^k - \bar{\theta}^k\|^2 = \|\Theta^k - 1(\bar{\theta}^k)^T\|^2_F.
\]

1) **Unbounded Operators**: If the operators \( F_i \) were bounded, i.e., there existed a constant \( D \) such that \( \|F_i(X, \theta)\| \leq D \) for all \( i \), \( X \), and \( \theta \), one can show that the norm of the consensus error is proportional to the step size \( \epsilon \), \( S^k \sim D \epsilon^2 \), which decays to zero as \( \epsilon \to 0 \) [29]. Based on the convergence of \( S^k \) the recursion on \( R^k \) can be easily established. Here, we make the less restrictive assumption that \( F_i \) is Lipschitz continuous (Assumption 2), allowing \( \|F_i(X, \theta)\| \) to be unbounded in general. In this case, one cannot immediately show that the consensus error converges to zero. Indeed, this error is proportional to the optimality error

\[
S^k \lesssim \delta S^{k-1} + O(\epsilon^2) R^{k-1}
\]

a.s., for some \( \delta < 1 \)
which can diverge unless $R_k^k$ decays to zero at a proper rate.

2) Markovian Samples: The data in our stochastic model are sampled from Markov processes. In other words, $(X_k^k)_k$ are dependent and have a shift in distribution from sample to sample. As a consequence, $F_k(X_k, \theta_k)$ is biased, i.e., $E_{\mu_k}[F_1(X_1, \theta_1)] \neq F_1(\theta_1)$, and automatic variance reduction does not follow by simply averaging across consecutive samples. In order to treat the Markovian data, one can utilize the recent development in analyzing (decentralized) SA under Markov samples (e.g., [13] and [16]). Specifically, one can take advantage of the geometric mixing time of the underlying Markov chain to handle the data dependence. Although the data are Markovian, their dependence is very weak at samples spaced out every $\tau_k$ step. However, this technique causes an extra “delay” consensus error in our analysis, i.e.,

$$\mathbb{E}[R_k^k] \leq \left(1 - \frac{\alpha_k}{2}\right)R_{k-1}^k + O(\epsilon)\mathbb{E}[S_{k-1}^k] + O(\epsilon)\mathbb{E}[S_{k-\tau_k}^k].$$

The delay term, $S_{k-\tau_k}^k$, implies the errors in the earlier iterates affect the current updates. This together with the coupling between consensus and optimality errors above complicates the analysis in Theorem 1. We note that the prior work [16] does not face these technical challenges since they consider a special linear case, i.e., $F_k$ are linear and the same for every agent. In that case, the consensus error is only proportional to $\epsilon$, similar to the existing literature [29]. In addition, there does not exist the delay term $S_{k-\tau_k}^k$ in the analysis in [16]. Without these coupling and delay terms, one can immediately see that the convergence of the consensus error implies the convergence of the optimality error to zero.

3) Novel Lyapunov Approach: As mentioned previously, one cannot study the convergence of consensus and optimality errors separately since these errors are coupled with each other. In addition, one needs to consider the impact of delays due to the Markovian samples. To handle this mixture of errors, a natural way is to analyze these errors simultaneously. Indeed, one of the main contributions in this article is to introduce a novel composite Lyapunov–Razumikhin function $V$ defined as

$$V^k = \mathbb{E}[R_k^k] + \mathbb{E}[S^k] + \mathbb{E}[S_{k-\tau_k}^k].$$

Our main motivation in using $V^k$ comes from the wide application of the Lyapunov–Razumikhin function in studying the stability of delayed ordinary differential equations [31]. By putting together the bounds of the optimality and consensus errors mentioned previously, we will show that $V^k$ decays to zero at a proper rate, which immediately implies the results in Theorem 1.

D. Time-Varying and Directed Communication Graphs

We note that with slight modifications, our algorithm and theoretical results extend to the case where the agents communicate according to a time-varying communication graph $G_k = (V, E_k)$. Essentially, we assume that the there exists a positive integer $B$ such that for any integer $l \geq 0$, the following graph is connected

$$(V, E_0 \cup E_1 \cup \ldots \cup E_{(l+1)\tau_k-1}).$$

This assumption is often considered in prior works that study consensus algorithms under a time-varying graph [15], [29], and guarantees the long-term connectivity though the network in any step can be disconnected. Let the consensus weight matrix $W_k$ be any double stochastic matrix that respects the graph $G_k$ (see Assumption 4). Our update rule in (2) becomes

$$\theta_k^k = \sum_{j \in \mathcal{N}_k(i)} W_k(i,j)\theta_j^k + \epsilon_k F_k(X_k^k, \theta_k^k)$$

where $\mathcal{N}_k(i)$ denotes the neighbors of the agent $i$ in iteration $k$. It has been observed in [29] that under a time-varying communication graph, the key structural quantity affecting the rate of convergence is

$$\eta = \min\{1 - 1/(2N^3), \min_{\ell \in \mathbb{Z}_+} \max_{-\ell \leq t \leq (\ell-1)B \leq (\ell-1)B - 1} \sigma_2(W_{\ell})\} \in [0, 1).$$

$\eta$ plays the role of $\sigma_2$ in the fixed communication graph setting. Our analysis can be adapted to show that with a time-varying communication graph, the modified update rule (15) produces iterates that converge linearly to a ball around the optimal solution $\theta^*$ under a constant step size; under a diminishing step size $\epsilon_k = O(1/k)$, the iterates converge exactly to $\theta^*$ with rate $O(\log^2(k)/k)$. These rates are the same as our results under a fixed communication graph, except for the constant $B$ capturing the connectivity of time-varying graphs.

Similarly, our analysis can be extended to the case where the communication graph is directed and strongly connected by leveraging the techniques in [32].

III. APPLICATIONS AND EXPERIMENTS

In this section, we illustrate our theoretical results by considering a number of numerical simulations on the performance of decentralized Markov chain gradient descent and decentralized Q-learning methods in solving problems in robust identification and reinforcement learning. We show that these two methods can be formulated as different variants of Algorithm 1.

A. Decentralized Markov Chain Gradient Descent

We consider a distributed optimization problem over a network of $N$ agents. Associated with agent $i$ is a function $\tilde{f}_i : \mathbb{R}^d \rightarrow \mathbb{R}$. The goal of the agents is to cooperatively solve

$$\min_{\theta \in \mathbb{R}^d} \sum_{i=1}^{N} \tilde{f}_i(\theta)$$

where $\tilde{f}_i(\theta) = \mathbb{E}_{X_i} f_i(X_i, \theta)$ for some $f_i : X_i \times \mathbb{R}^d \rightarrow \mathbb{R}$ and $X_i$ is a random variable taking values over the space $X_i$ with the stationary distribution $\mu_i$. DSGD is one of the most popular methods in the literature for solving this problem, where each agent $i$ maintains $\theta_i$, an estimate of the solution $\theta^*$ of (16), and iteratively performs

$$\theta_i^{k+1} = \sum_{j \in \mathcal{N}_i} W_{ij} \theta_j^k + \epsilon_k \nabla f_i(X_i^k, \theta_i^k).$$

Most of the existing results about the convergence of (17) are established under the conditions that $X_i^k$ are i.i.d. across time and $i$. However, in many problems such as optimization under a linear dynamic system, $\{X_i^k\}$, are sampled from Markov chains. See [10] and [33] for a discussion on more applications where the data are naturally Markovian.

As a concrete example, we consider the distributed variant of the robust identification problem [34]. In this problem, each agent $i$ is associated with a system as shown in Fig. 1. The system parameters are $A_i \in \mathbb{R}^{d \times d}$, which is different across agents, and $u \in \mathbb{R}^d$, the same across agents. The observations (samples) are $X_i^k = (X_i^k(1), X_i^k(2)) \in \mathbb{R}^d \times \mathbb{R}$, generated by the following
that solves is the first basis vector. We then generate − is the quadratic lost function is drawn uniformly from \( N = 2 \} \) e \( \{ \nabla System ID sample generation. \)

\[
y(2) = \mu_e \{ y \in [y_e] \} \]

is \( u_0, \in = 5 \) 0 \( \mu = f_1 \geq u = 2 \) 0 (19) (1) = only connect to agents 2 and \( \langle u, X_i(1) \rangle \) + \( X_i(1) \) + \( y_i \) \( k \) \( d \times d \) \( \mu \) i.i.d. \( -F_i = \nabla_\theta f_i \) is strongly monotone if \( \bar{F}_i \) is positive definite. Under the strong monotonicity of \( -F_i \), our theoretical results in Theorem 1 imply the same results for the finite-time perfomance of the decentralized Markov SGD. Our rates recover the best known result of decentralized Markov SGD while relying on weaker assumptions [10].

To illustrate our theoretical results, we experimentally solve this problem where we choose \( N = 50 \) and \( d = 20 \). To generate data, we choose \( X_i \in \mathbb{R}^{d \times d} \) to be a subdiagonal matrix where for all \( 2 \leq m \leq d \), \( A, [m, m-1] \) is drawn uniformly from \([.8,.99]\), and \( u \) to be a random vector drawn uniformly from the d-dimensional unit \( \ell_2 \)-ball. We sample scalars \( u_i \) and \( y_i \) i.i.d. from the standard normal distribution \( N(0,1) \), and make \( U_i = u_i \epsilon_1 \), where \( \epsilon_1 \) is the first basis vector. We then generate the samples by making \( X_i(1) = 0 \) and following (18) for \( k \geq 1 \).

In this experiment, we use a line graph \( G \), i.e., the agent \( i \) connects to agents \( i - 1 \) and \( i + 1 \) for all \( i = 2, 3, \ldots, N - 1 \), and agents 1 and \( N \) only connect to agents 2 and \( N - 1 \), respectively. In addition, we generate \( W \) using the lazy metropolis method [38].

In Fig. 2, we plot the decay of the errors under a constant step size \( \epsilon = 5e - 4 \) (top row) and a diminishing step size with \( \epsilon = 3e - 2 \) (bottom row). In the case of the constant step size, we see that the optimality error roughly converges linearly after the initial 2000 iterations, which matches the theory. The decay of the consensus error is much faster, which is a common result and observation in the literature of consensus algorithms [29]. In the case of diminishing step size, both the optimality error and the consensus error decays roughly with rate \( O(\log^2(k)/k) \), again matching our theoretical result.

**B. Decentralized Q Learning**

We consider multitask reinforcement learning problems over multiagent systems, where there are \( N \) different tasks and \( N \) agents [19]. Each task, modeled by a Markov decision process, is assigned to one agent. The goal of the agents is to cooperatively find a single policy, a mapping from state spaces to the action spaces, that simultaneously solves all the tasks.

For solving this problem, we consider a decentralized variant of Q-learning method under linear function approximation [14]. When the tasks share the same dynamic, this problem is equivalent to find a \( \theta^* \) that solves

\[
\sum_{i=1}^{N} \mathbb{E}_{X \sim \mu_i} [G_i(\theta^*, X)] = 0
\]

for some nonlinear operator \( G_i \) that depends on the basis of the linear subspace and the reward functions at different tasks. For more details, we refer the readers to [14], which discusses the problem in the single agent case. Under a proper assumption on \( \mu_i \) and \( G_i \) that guarantees sufficient state visitation, one can show that the operator \( -G_i(\theta) = -\mathbb{E}_{X \sim \mu_i} [G_i(\theta, X)] \) is strongly monotone. With \( F_i = -G_i \), our results in Theorem 1 can be applied to derive the rates of decentralized Q-learning, which is unknown in the existing literature.

To illustrate for the performance of decentralized Q-learning, we consider a GridWorld problem consisting of multiple mazes (environments) with size \( 10 \times 10 \) (see the first plot of Fig. 3). In each environment, the agent is placed in a grid of cells, where every cell has one of three labels: goal, obstacle, or...
Lemma 1 is a simple consequence of the Lipschitz condition of \( F_i(X, \cdot) \), and states that the norm of the operator \( F_i(X, \theta) \) or \( \tilde{F}_i(\theta) \) can be bounded affinely in \( \theta \).

**Lemma 2:** For all \( k \geq \tau_k \), we have
\[
\| \tilde{\theta}^k - \tilde{\theta}^{k-\tau_k} \| \leq 3\epsilon_{k-\tau_k} B \tau_k \left( \| \tilde{\theta}^k \| + \frac{1}{N} \sqrt{S^{k-\tau_k}} + 1 \right)
\]
\[
\leq \frac{1}{2N} \left( \| \tilde{\theta}^k \| + \sqrt{S^{k-\tau_k}} + 1 \right). \tag{20}
\]

2) For all \( k \leq \tau_k \), we have
\[
\| \tilde{\theta}^k - \tilde{\theta}^0 \| \leq \frac{1}{3N} \| \tilde{\theta}^0 \| + \frac{1}{3N} \sqrt{S^0} + \frac{1}{3N} \]
\[
\text{where the constant}
\]
\[
C_0 = 32B^2 \left( \| \tilde{\theta}^0 \|^2 + 1 \right).
\]

**Lemma 3:** We have for all \( k \geq 0 \)
\[
S^k \leq \frac{1 + \sigma^2}{2} S^{k-1} + \frac{32B^2 N}{1 - \sigma^2} R^{k-1} + \frac{NC_0}{1 - \sigma^2} \]
\[
\text{where the constant}
\]
\[
C_0 = 32B^2 \left( \| \tilde{\theta}^0 \|^2 + 1 \right).
\]

**Lemma 4:** We have for all \( k \geq \tau_k \)
\[
E[R^{k+1}] \leq \left( 1 - \frac{\alpha\epsilon_{k-\tau_k}}{2} \right) E[R^k] + NC_1 \epsilon_{k-\tau_k} \tau_k
\]
\[
\quad + NC_2 \epsilon_{k} E[S^k] + NC_2 \epsilon_{k} E[S^{k-\tau_k}]
\]
\[
\text{where} \ C_1 \text{ and } C_2 \text{ are defined in (9).}
\]

Lemmas 3 and 4 characterize the decay of the consensus error and the optimality error, respectively. For conciseness, the proofs of the lemmas are deferred to the appendix.

**B. Proof of Theorem 1**

\( V^k = E[R^k] + E[S^k] + E[S^{k-\tau_k}] \) is the sum of three components. Bounding \( V^k \) requires characterizing the growth of each of the three terms, which are all coupled with each other.

By Lemmas 3 and 4, we can recursively bound the consensus error
\[
E[S^k] \leq \frac{1 + \sigma^2}{2} E[S^{k-1}]
\]
\[
\quad + \frac{32B^2 N}{1 - \sigma^2} E[R^{k-1}] + \frac{NC_0}{1 - \sigma^2} \]
\[
\text{and the optimality error}
\]
\[
E[R^k] \leq \left( 1 - \frac{\alpha\epsilon_{k-1}}{2} \right) E[R^{k-1}] + NC_1 \epsilon_{k-1} \epsilon_{k-\tau_{k-1}-1} \tau_{k-1}
\]
\[
\quad + NC_2 \epsilon_{k-1} E[S^{k-1}] + NC_2 \epsilon_{k-1} E[S^{k-\tau_{k-1}-1}]
\]
\[
\leq \left( 1 - \frac{\alpha\epsilon_{k-1}}{2} \right) E[R^{k-1}] + NC_1 \epsilon_{k-1} \epsilon_{k-\tau_{k-1}} \tau_{k-1}
\]
\[
\quad + NC_2 \epsilon_{k-1} E[S^{k-1}] + NC_2 \epsilon_{k-1} E[S^{k-\tau_{k-1}-1}]
\]
\[
\text{where the second inequality follows from} \ \tau_{k-1} \leq \tau_k \text{ and} \ \epsilon_{k-\tau_{k-1}-1} \leq \epsilon_{k-\tau_k}.
\]

Let \( i^* \) denote index of the agent with the largest \( \rho_i \), i.e., \( \rho_{i^*} = \rho \). For any \( k \geq \frac{1}{\rho^2} \), (3), the definition of the mixing time, and the fact that the max operator is nonexpansive imply
\[
\tau_k - \tau_{k-1} \leq \log_{\rho} \left( \frac{\epsilon_k}{m_i^*} \right) - \log_{\rho} \left( \frac{\epsilon_{k-1}}{m_i^*} \right)
\]
\[ \log \rho \left( \frac{\epsilon_k}{\epsilon_{k-1}} \right) = \log \rho \left( \frac{k + 1}{k} \right) \leq 1. \]

Since \( \tau_k, \tau_{k-1} \) are both integers and \( \tau_k \geq \tau_{k-1} \), we have either \( \tau_k = \tau_{k-1} + 1 \) or \( \tau_k = \tau_{k-1} + 1 \).

When \( \tau_k = \tau_{k-1} + 1 \), combining (23) and (24) implies \( \forall k \geq \tau_k \)

\[ V^k = \mathbb{E}[R^k] + \mathbb{E}[S^k] + \epsilon_{k-1} \tau_k \]

\[ \leq \left( 1 - \frac{\alpha_k - 1}{2} \right) \mathbb{E}[R^{k-1}] + NC_1 \epsilon_{k-1} \mathbb{E}[S^{k-1}] \]

\[ + NC_2 \epsilon_{k-1} \mathbb{E}[S^k] + \frac{1 + \sigma^2_{\tau_k}}{2} \mathbb{E}[S^k] + \frac{32N^2 \epsilon_{k-1}^2}{1 - \sigma^2_k} \mathbb{E}[R^{k-1}] \]

\[ + \frac{N C_2}{1 - \sigma^2_k} \epsilon_{k-1}^2 + \frac{1 + \sigma^2_{\tau_k}}{2} \mathbb{E}[S^{k-1}] \]

\[ \leq \left( 1 - \frac{\alpha_k - 1}{4} + \frac{32N^2 \epsilon_{k-1}^2}{1 - \sigma^2_k} \right) \mathbb{E}[R^{k-1}] \]

\[ + \frac{N C_2}{1 - \sigma^2_k} \epsilon_{k-1}^2 + \frac{1 + \sigma^2_{\tau_k}}{2} \mathbb{E}[S^{k-1}] \]

\[ \leq \left( 1 - \frac{\alpha_k - 1}{4} + \frac{32N^2 \epsilon_{k-1}^2}{1 - \sigma^2_k} \right) \mathbb{E}[R^{k-1}] \]

\[ + \frac{N C_2}{1 - \sigma^2_k} \epsilon_{k-1}^2 + \frac{1 + \sigma^2_{\tau_k}}{2} \mathbb{E}[S^{k-1}] \]

\[ \leq \frac{1}{1 - \sigma^2_k} \mathbb{E}[S^k] \]

\[ \leq \frac{1}{1 - \sigma^2_k} \mathbb{E}[S^k] \]

Note that by choosing the step size small enough, the terms quadratic in the step size is dominated by terms linear in the step size. Specifically, if the step size sequence satisfies

\[ \epsilon_k \leq \min \left\{ \frac{c_3 \alpha(1 - \sigma^2_\tau)}{768N^2 \epsilon^2_\sigma}, \frac{1 - \sigma^2_\tau}{4} + \frac{12N^2 \epsilon^2_\sigma}{c_3(1 - \sigma^2_\sigma)} + 2NC_2 \right\} \quad \forall t \geq 0 \]

which is guaranteed by the conditions of Theorem 1, then we can further simplify the aforementioned inequality with the help of (6) and the fact \( \epsilon_{k-1} \leq \frac{c_3}{c_0} \) for all \( k \geq \tau_k \), which is a result of (6)

\[ V^k \leq \left( 1 - \frac{\alpha_k - 1}{8} \right) \mathbb{E}[R^{k-1}] \]

\[ + \frac{N (C_1 + \frac{2C_0}{1 - \sigma^2_k} + \frac{192N^2 \epsilon^2_\sigma}{1 - \sigma^2_k} \left( \mathbb{E}[\|\theta^*\|^2] + 2 \right) + \frac{1 + \sigma^2_{\tau_k}}{2} \mathbb{E}[S^{k-1}] \]
where the last inequality uses $\epsilon_{k-1} \leq \frac{C_2(1 - \sigma^2)}{2B^2}$.

Defining

$$
\lambda_k = \begin{cases} 
1 - \frac{\alpha \epsilon_k}{8}, & \text{if } \tau_{k+1} = \tau_k \\
1 + NC_2 \epsilon_k, & \text{if } \tau_{k+1} = \tau_k + 1
\end{cases}
$$

and recursively applying the inequalities (26) and (27), we get for all $k \geq \tau_k$

$$
V_k \leq V^{\tau_k} \prod_{t = \tau_k}^{k-1} \lambda_t + \frac{13NC_1}{1 - \sigma^2} \sum_{t = \tau_k}^{k-1} \epsilon^2 \tau (e) \prod_{t = \tau_k}^{k-1} \left( 1 - \frac{\alpha \epsilon_k}{8} \right)^{k - t - 1} \lambda_t.
$$

The aforementioned equation characterizes the general convergence behavior under either choice of the step sizes of Theorem 1. We now specialize to each step size sequence.

1) Constant Step Size: With a constant step size $\epsilon_k = \epsilon$ (hence, $\tau_k = \tau_{k-1} = \tau(e)$ for all $k$), (28) implies

$$
V^k \leq V^{\tau(e)} \prod_{t = \tau(e)}^{k-1} \left( 1 - \frac{\alpha \epsilon}{8} \right)^{k - t} + \frac{13NC_1}{1 - \sigma^2} \sum_{t = \tau(e)}^{k-1} \epsilon^2 \tau (e) \prod_{t = \tau(e)}^{k-1} \left( 1 - \frac{\alpha \epsilon}{8} \right)^{k - t - 1}
$$

$$
\leq V^{\tau(e)} \left( 1 - \frac{\alpha \epsilon}{8} \right)^{k - \tau(e)} + \frac{13NC_1}{1 - \sigma^2} \epsilon^2 \tau (e) \prod_{t = \tau(e)}^{k-1} \left( 1 - \frac{\alpha \epsilon}{8} \right)^{k - t - 1}
$$

$$
\leq V^{\tau(e)} \left( 1 - \frac{\alpha \epsilon}{8} \right)^{k - \tau(e)} + \frac{104NC_1}{\alpha(1 - \sigma^2)} \epsilon^2 \tau (e).
$$

To bound $V^{\tau(e)}$ in terms of the initial conditions, we first bound $S^{\tau(e)}$. Applying Lemma 3 recursively, we have

$$
S^{\tau(e)} \leq \frac{1 + \sigma^2}{2} S^{\tau(e)-1} + \frac{32N \epsilon^2 B^2}{1 - \sigma^2} R^{\tau(e)-1} + \frac{NC_0}{1 - \sigma^2} \epsilon^2
$$

$$
\leq S^{\tau(e)-1} + \frac{32N \epsilon^2 B^2}{1 - \sigma^2} R^{\tau(e)-1} + \frac{NC_0}{1 - \sigma^2} \epsilon^2
$$

$$
\leq S^0 + N \epsilon^2 \sum_{t = \tau(e)}^{\tau(e)-1} \left( \frac{32B^2}{1 - \sigma^2} R^t + \frac{C_0}{1 - \sigma^2} \right).
$$

By (21) of Lemma 2 and the Cauchy–Schwarz inequality

$$
R^t \leq 2(\|\bar{\theta} - \tilde{\theta}\|^2 + R^0)
$$

$$
\leq 2 \left( \frac{1}{3N} \|\tilde{\theta}\|^2 + \frac{1}{3N} \sqrt{S^0} + \frac{1}{3N} \right)^2 + 2R^0
$$

$$
\leq \frac{2}{3N} (\|\tilde{\theta}\|^2 + S^0 + 1) + 2R^0.
$$

Then, (29) becomes

$$
S^{\tau(e)} \leq S^0 + N \epsilon^2 \sum_{t = 0}^{\tau(e)-1} \left( \frac{64B^2}{3N^2(1 - \sigma^2)} (\|\tilde{\theta}\|^2 + S^0 + 1) + \frac{64B^2}{1 - \sigma^2} R^0 + \frac{C_0}{1 - \sigma^2} \right)
$$

$$
\leq S^0 + N \epsilon^2 \tau (e) \left( \frac{128B^2}{3N^2(1 - \sigma^2)} (R^0 + \|\tilde{\theta}\|^2) + \frac{64B^2}{1 - \sigma^2} R^0 + \frac{C_0}{1 - \sigma^2} \right)
$$

$$
\leq \left( 1 + \frac{16}{27N^3(1 - \sigma^2)} \right) S^0 + \frac{16B^2}{9N(1 - \sigma^2)} R^0
$$

$$
+ \frac{C_0}{9NB^2(1 - \sigma^2)}
$$

(30)

where the last inequality combines the terms with the relations $\epsilon \tau (e) \leq \frac{1}{4N}$ and $\tau (e) \geq 1$.

We then bound $R^{\tau(e)}$ as

$$
R^{\tau(e)} \leq 2(\|\tilde{\theta} - \tilde{\theta}\|^2 + R^0)
$$

$$
\leq \frac{2}{3N} (\|\tilde{\theta}\|^2 + S^0 + 1)^2 + 2R^0
$$

$$
\leq \left( 4 + \frac{2}{3N} \right) (\|\tilde{\theta}\|^2 + \frac{2}{3N^2} S^0 + 4\|\tilde{\theta}\|^2 + \frac{2}{3N^2} )
$$

(31)

where the second inequality again comes from (21) of Lemma 2. Then, we can plug (30) and (31) into $V^{\tau(e)}$ as

$$
V^{\tau(e)} = \mathbb{E}[S^{\tau(e)}] + \mathbb{E}[S^{\tau(e)}] + S^0
$$

$$
\leq \left( 4 + \frac{2}{3N^2} \right) (\|\tilde{\theta}\|^2 + \frac{2}{3N^2} S^0 + 4\|\tilde{\theta}\|^2 + \frac{2}{3N^2} )
$$

$$
+ \left( 1 + \frac{16}{27N^3(1 - \sigma^2)} \right) S^0
$$

$$
+ \frac{16}{9N(1 - \sigma^2)} R^0 + \frac{C_0}{9NB^2(1 - \sigma^2)} + S^0
$$

$$
\leq \left( 2 + \frac{16}{27N^3(1 - \sigma^2)} \right) S^0 + \frac{2}{3N^2}
$$

$$
+ \left( 4 + \frac{2}{3N^2} + \frac{16}{9N(1 - \sigma^2)} \right) \|\tilde{\theta}\|^2
$$

$$
+ \frac{16}{9N(1 - \sigma^2)} \|\tilde{\theta}\|^2 + \frac{C_0}{9NB^2(1 - \sigma^2)}
$$

$$
\leq \frac{4}{1 - \sigma^2} S^0 + \frac{7}{1 - \sigma^2} \|\tilde{\theta}\|^2 + \frac{C_1}{1 - \sigma^2}
$$

(32)

where the last inequality follows from simplification using $N \geq 1$ and the definition of $C_0$ in (22).

Therefore, for any $i = 1, 2, \ldots, N$ and any $k > \tau(e)$

$$
\mathbb{E}[\|\theta^i_k - \theta^i_k\|^2] = \mathbb{E}[\|\theta^i_k - \theta^i_k + (\tilde{\theta} - \bar{\theta})\|^2]
$$

$$
\leq 2\mathbb{E}[R^k] + 2\mathbb{E}[\|\theta^i_k - \theta^i_k\|^2] \leq 2\mathbb{E}[R^k] + 2\mathbb{E}[S^k] \leq 2V^k
$$

$$
\leq \left( 1 - \frac{\alpha \epsilon}{8} \right)^{k - \tau(e)} V^{\tau(e)} + \frac{208NC_1}{\alpha(1 - \sigma^2)} \epsilon \tau (e)
$$

$$
\leq \left( 1 - \frac{\alpha \epsilon}{8} \right)^{k - \tau(e)} \left( \frac{4}{1 - \sigma^2} S^0 + \frac{7}{1 - \sigma^2} \|\tilde{\theta}\|^2 + \frac{C_1}{1 - \sigma^2} \right)
$$

+ \frac{208N \beta C_1}{\alpha(1 - \sigma^2)} \epsilon \log \left( \frac{1}{\epsilon} \right).

2) Diminishing Step Size: Since $\epsilon_k$ is a nonincreasing function in $k$ and $\tau_k = \tau_{k-1} + 1$ occurs no more than $\tau_k$ times when $t$ increases from $\tau_k$ to $k$, (28) implies

$$
V^k \leq \prod_{t = \tau_k}^{k-1} \lambda_t + \frac{13NC_1}{1 - \sigma^2} \sum_{t = \tau_k}^{k-1} \epsilon^2 \tau t - \tau_k + 1 \prod_{t = t+1}^{k-1} \lambda_t
$$

(33)
\[ V^{\tau_k} \left( \prod_{t=\tau_k}^{\tau_{k+1}} (1 + NC_2 \epsilon_t) \right) \left( \prod_{t=\tau_k}^{\tau_{k+1}} \left( 1 - \frac{\alpha \epsilon_t}{8} \right) \right) \]
\[ + \frac{13NC_1}{1 - \sigma^2} \sum_{t=\tau_k}^{\tau_{k+1}} \epsilon_{t-\tau_k-1} \epsilon_t \left( \prod_{t=\tau_k}^{\tau_{k+1}} (1 + NC_2 \epsilon_t) \right) \times \left( \prod_{t=\tau_k}^{\tau_{k+1}} \left( 1 - \frac{\alpha \epsilon_t}{8} \right) \right). \]

(33)

Since \( 1 + x \leq \exp(x) \) for any scalar \( x \), we have \( \forall \ell \geq \tau_k \)

\[ \left( \prod_{t=\ell}^{\ell+\tau_k} (1 + NC_2 \epsilon_t) \right) \left( \prod_{t=\ell+\tau_k}^{\ell+\tau_{k+1}} \left( 1 - \frac{\alpha \epsilon_t}{8} \right) \right) \leq \exp(NGC_2 \epsilon_{\tau_k}) \left( \exp \left( - \sum_{t=\ell+\tau_k}^{\ell+\tau_{k+1}} \frac{\alpha \epsilon_t}{8} \right) \right) \]

\[ \leq 2 \exp \left( - \sum_{t=\ell+\tau_k}^{\ell+\tau_{k+1}} \frac{\alpha \epsilon_t}{8(t+1)} \right) \]

where the third inequality uses \( NC_2 \epsilon_{\tau_k} \leq 1/2 \leq \log(2) \) afforded by the step size condition. Note that

\[ \sum_{t=k_1}^{k_2} \frac{1}{t+1} \geq \int_{k_1}^{k_2+1} \frac{1}{t+1} dt = \log \left( \frac{k_2 + 2}{k_1 + 1} \right) \]

which simplifies the aforementioned inequality:

\[ \left( \prod_{t=\ell}^{\ell+\tau_k} (1 + NC_2 \epsilon_t) \right) \left( \prod_{t=\ell+\tau_k}^{\ell+\tau_{k+1}} \left( 1 - \frac{\alpha \epsilon_t}{8} \right) \right) \]

\[ \leq 2 \exp \left( - \sum_{t=\ell+\tau_k}^{\ell+\tau_{k+1}} \frac{\alpha \epsilon_t}{8(t+1)} \right) \leq \frac{2(\ell + \tau_k + 2)}{k + 1} \]

if \( \alpha \geq 8 \). Using this bound in (33), we have

\[ V^k \leq \frac{4(\tau_k + 1)}{k + 1} V^{\tau_k} + \frac{13NC_1}{1 - \sigma^2} \sum_{t=\tau_k}^{\tau_{k+1}} \epsilon_{t-\tau_k-1} \epsilon_t \frac{2(\ell + \tau_k + 1)}{k + 1} \]

\[ \leq \frac{4(\tau_k + 1)}{k + 1} V^{\tau_k} + \frac{52NC_1}{1 - \sigma^2} \tau_k - \tau - 1 + \frac{1}{k + 1} \]

\[ \leq \frac{4(\tau_k + 1)}{k + 1} V^{\tau_k} + \frac{52NC_1 \tau_k - \tau - 1}{c_k^2(1 - \sigma^2)} \sum_{t=\tau_k}^{\tau_{k+1}} \frac{1}{\ell^2} \]

\[ \leq \frac{4(\tau_k + 1)}{k + 1} V^{\tau_k} + \frac{208NC_1 \tau_k - \tau - 1}{c_k^2(1 - \sigma^2)} \sum_{t=\tau_k}^{\tau_{k+1}} \frac{1}{\ell^2} \]

where the third inequality uses \( \epsilon_{k-\tau_k} \leq \frac{\epsilon}{\tau_k} \) for all \( k \geq \tau_k \), which is a result of (6), and the fourth inequality follows from \( \frac{t+1}{t^2} \leq \frac{1}{t^2} \) for all \( t \geq 1 \). Applying

\[
\sum_{t=k_1}^{k_2} \frac{1}{t+1} \leq \int_{k_1-1}^{k_2} \frac{1}{t+1} dt = \log \left( \frac{k_2 + 1}{k_1} \right)
\]

in (34), we have

\[
V^k \leq \frac{4(\tau_k + 1)}{k + 1} V^{\tau_k} + \frac{208NC_1 \tau_k - \tau - 1}{c_k^2(1 - \sigma^2)} \sum_{t=\tau_k}^{\tau_{k+1}} \frac{1}{\ell^2} \]

in (34), we have

\[
V^k \leq \frac{4(\tau_k + 1)}{k + 1} V^{\tau_k} + \frac{208NC_1 \tau_k - \tau - 1}{c_k^2(1 - \sigma^2)} \sum_{t=\tau_k}^{\tau_{k+1}} \frac{1}{\ell^2} \]

We can bound \( V^{\tau_k} \) by the initial conditions the same way as in the case of a constant step size and get

\[
V^{\tau_k} \leq \frac{4}{1 - \sigma^2} + \frac{7}{1 - \sigma^2} \parallel \tilde{\theta}^0 \parallel^2 + \frac{C_1}{1 - \sigma^2}.
\]

Then, similar to (32), for any \( i \in \{1, 2, \ldots, N\} \) and any \( k > \tau_k \)

\[
\mathbb{E} \left[ \parallel \tilde{\theta}^k - \theta^* \parallel^2 \right] \leq 2V^k
\]

\[
\leq \frac{4(\beta \log(k + 1) - \log(\epsilon)) + 1}{(1 - \sigma^2)} \parallel \tilde{\theta}^0 \parallel^2 + \frac{C_1}{1 - \sigma^2}.
\]

The same argument extends easily to \( \parallel \tilde{F}_i(\tilde{\theta}) \parallel \).

\[
\| \tilde{F}_i(X_i, \tilde{\theta}) \| - \| F_i(X_i, 0) \|
\]

\[
\leq \| F_i(X_i, \tilde{\theta}) - F_i(X_i, 0) \| \leq L \| \tilde{\theta} \|.
\]

Rearranging terms, we have for all \( i \) and \( X_i \)

\[
\| F_i(X_i, \tilde{\theta}) \| \leq L \| \tilde{\theta} \| + \| F_i(X_i, 0) \|
\]

\[
\leq B \| \tilde{\theta} \| + B \leq B(\| \tilde{\theta} \| + 1).
\]

**APPENDIX A**

**PROOF OF TECHNICAL LEMMAS**

**A. Proof of Lemma 1**

Using Assumption 2

\[ \| F_i(X_i, \tilde{\theta}) \| - \| F_i(X_i, 0) \| \]

\[ \leq \| F_i(X_i, \tilde{\theta}) - F_i(X_i, 0) \| \leq L \| \tilde{\theta} \|. \]

Rearranging terms, we have for all \( i \) and \( X_i \)

\[ \| F_i(X_i, \theta) \| \leq L \| \tilde{\theta} \| + \| F_i(X_i, 0) \| \]

\[ \leq B \| \tilde{\theta} \| + B \leq B(\| \tilde{\theta} \| + 1). \]

The same argument extends easily to \( \| \tilde{F}_i(\tilde{\theta}) \| \).

**B. Proof of Lemma 2**

As a result of Lemma 1

\[ \| \tilde{\theta}^{k+1} \| - \| \tilde{\theta}^k \| \leq \| \tilde{\theta}^{k+1} - \tilde{\theta}^k \| \]

\[ = \| \sum_{i=1}^{N} F_i(X_i^k, \tilde{\theta}^k) \| \]

\[ \leq \| \sum_{i=1}^{N} B(\| \tilde{\theta}_i^k \| + 1) \|
\]

\[ \leq \sum_{i=1}^{N} \| \tilde{\theta}_i^k \| + \| \tilde{\theta}_i^k - \tilde{\theta}_i^k \| + 1 \]

\[ \leq \epsilon_k B(\| \tilde{\theta}^k \| + \sqrt{S^k} + 1). \]
The last inequality of (35) follows from the fact that 
\[ ||\theta^k - \bar{\theta}|| \leq \sqrt{S^k} \]
which is a simple consequence of the definition of \( S^k \) in (14). 
Rearranging the terms in (35) leads to 
\[ ||\bar{\theta}^{k+1}|| \leq (1 + c_k B)||\bar{\theta}|| + c_k B(\sqrt{S^k} + 1). \]
The matrix form of the update rule (2) is 
\[ \Theta^{k+1} = W\Theta^k + e_k F(X^k, \Theta^k). \]  
(36)
The aforementioned equation implies that the update for \( \bar{\theta} \) observes 
\[ \bar{\theta}^{k+1} = \frac{1}{N}(\Theta^k)\top + \frac{c_k}{N} F(X^k, \Theta^k)\top 1. \]  
(37)
Using the update rule of \( \Theta^{k+1} \) and \( \bar{\theta}^{k+1} \)
\[ \Theta^{k+1} = W\Theta^k - \frac{1}{N} 11\top \Theta^k \]
\[ + c_k \left( F(X^k, \Theta^k) - \frac{1}{N} 11\top F(X^k, \Theta^k) \right) \]
\[ = (W - \frac{1}{N} 11\top)\Theta^k - (W - \frac{1}{N} 11\top 1)\bar{\theta}^{k}\top \]
\[ + c_k \left( F(X^k, \Theta^k) - \frac{1}{N} 11\top F(X^k, \Theta^k) \right) \]
\[ = (W - \frac{1}{N} 11\top)\left( \Theta^k - \frac{1}{N}(\bar{\theta}^{k})\top \right) \]
\[ + c_k \left( I - \frac{1}{N} 11\top \right) F(X^k, \Theta^k). \]  
(38)
Note that for any matrix \( Y \in \mathbb{R}^{N \times N} \), \( Z \in \mathbb{R}^{N \times d} \), we have \( ||YZ||_F \leq ||Y||_2 ||Z||_F \), so that 
\[ ||YZ||_F^2 = \sum_{i=1}^{N} ||Yz_i||^2 \leq \sum_{i=1}^{N} ||Z||_2 ||z_i||^2 \]
\[ \leq ||Y||_2^2 \sum_{i=1}^{N} ||z_i||^2 = ||Y||_2^2 ||Z||_F^2, \]
where \( z_i \) denotes the \( i \)th column of \( Z \). This inequality implies 
\[ ||\Theta^{k+1} - \frac{1}{N}(\bar{\theta}^{k})\top ||_F \leq \left[ ||W - \frac{1}{N} 11\top \right] \sqrt{S^k} + c_k \left[ \frac{1}{N} 11\top \right] \left[ F(X^k, \Theta^k) \right]_F \]
\[ \leq \sqrt{S^k} + c_k \sum_{i=1}^{N} ||F_i(X^k, \theta^k_i)|| \]
\[ \leq \sqrt{S^k} + c_k \sum_{i=1}^{N} ||\theta^k_i|| + 1 \]
\[ \leq \sqrt{S^k} + c_k B \sum_{i=1}^{N} (||\theta^k_i - \bar{\theta}|| + ||\bar{\theta}|| + 1) \]
\[ \leq (1 + c_k N B) \sqrt{S^k} + c_k N B ||\bar{\theta}|| + c_k N B \]
where the first inequality follows from \( ||W - \frac{1}{N} 11\top ||_F \leq 1 \) and 
\[ ||I - \frac{1}{N} 11\top||_F \leq 1. \]

Define \( y^k = ||\bar{\theta}|| + \sqrt{S^k} \). Using the aforementioned two inequalities 
\[ y^{k+1} = ||\bar{\theta}^{k+1}|| + \sqrt{S^{k+1}} \]
\[ \leq (1 + c_k B)||\bar{\theta}|| + c_k B(\sqrt{S^k} + 1) \]
\[ + (1 + c_k N B)\sqrt{S^k} + c_k N B ||\bar{\theta}|| + c_k N B \]
\[ \leq (1 + 2 c_k N B)||\bar{\theta}|| + (1 + 2 c_k N B)\sqrt{S^k} + 2 c_k N B \]
\[ = (1 + 2N c_k B)y^k + 2c_k N B. \]
First, we show the bound for \( k \geq \tau_k \). For \( t' \in [k - \tau_k, k] \), 
\[ y^{t'} \leq (1 + 2c_{k-\tau_k} N B)^{t'-k+\tau_k} y^{k-\tau_k} \]
\[ + 2c_{k-\tau_k} N B \sum_{u=k-\tau_k}^{t'-1} (1 + 2c_{k-\tau_k} N B)^{u-t'} \]
\[ \leq (1 + 2c_{k-\tau_k} N B)^{\tau_k} y^{k-\tau_k} \]
\[ + 2c_{k-\tau_k} N B \tau_k (1 + 2c_{k-\tau_k} N B)^{\tau_k} \]
\[ \leq 2y^{k-\tau_k} + 4c_{k-\tau_k} N B \tau_k \]
(39)
where we have used the fact that \( 1 + x \leq e^x \) for \( x \geq 0 \) and the relationship \( e^{2c_{k-\tau_k} N B \tau_k} \leq 2 \), which comes from the condition on the step size \( 2c_{k-\tau_k} N B \tau_k \leq \frac{1}{4} \log(2) \). Then 
\[ ||\bar{\theta}^k - \bar{\theta}^{k-\tau_k}|| \leq \sum_{t'=k-\tau_k}^{k-1} (||\bar{\theta}^{t'}|| + \sqrt{S^{t'}} + 1) \]
\[ \leq c_{k-\tau_k} B \sum_{t'=k-\tau_k}^{k-1} (||\bar{\theta}^{t'}|| + \sqrt{S^{t'}} + 1) \]
\[ \leq c_{k-\tau_k} B \sum_{t'=k-\tau_k}^{k-1} (y^{t'} + 1) \]
\[ \leq c_{k-\tau_k} B \sum_{t'=k-\tau_k}^{k-1} (2y^{k-\tau_k} + 4c_{k-\tau_k} N B \tau_k + 1) \]
\[ \leq c_{k-\tau_k} B \tau_k (2||\bar{\theta}^{k-\tau_k}|| + 2\sqrt{S^{k-\tau_k}} + 2) \]
\[ \leq 2c_{k-\tau_k} B \tau_k |\bar{\theta}^{k-\tau_k} - \bar{\theta}^{k-\tau_k}| + 2c_{k-\tau_k} B \tau_k ||\bar{\theta}|| \]
\[ + 2c_{k-\tau_k} B \tau_k \sqrt{S^{k-\tau_k}} + 2c_{k-\tau_k} B \tau_k \]
(40)
where the second inequality is from (35), and the second last inequality comes from the step size rule \( c_{k-\tau_k} \tau_k \leq \frac{1}{4} \log(2) \). 
Rearranging terms and again using the condition \( c_{k-\tau_k} \tau_k \leq \frac{1}{4} \log(2) \), we have 
\[ ||\bar{\theta}^k - \bar{\theta}^{k-\tau_k}|| \]
\[ \leq 3c_{k-\tau_k} B \tau_k ||\bar{\theta}|| + 3c_{k-\tau_k} B \tau_k \sqrt{S^{k-\tau_k}} + 3c_{k-\tau_k} B \tau_k \]
\[ \leq \frac{1}{2N} ||\bar{\theta}|| + \frac{1}{2N} \sqrt{S^{k-\tau_k}} + \frac{1}{2N}. \]
Now it remains to show the bound on \( ||\hat{\theta}^k - \bar{\theta}|| \) for \( k \leq \tau_k \). Similar to (39), for \( t' \in [0, k] \), we have
\[
y^{t'} \leq 2y^0 + 4\epsilon_k N B \tau_k = 2y^0 + 4\epsilon_k N B \tau_k
\]
which implies for all \( k \leq \tau_k \)
\[
\|\hat{\theta}^k - \bar{\theta}^0\| \leq \sum_{t'=0}^{k-1} \|\hat{\theta}^{t'+1} - \bar{\theta}^{t'}\|
\]
\[
\leq \epsilon B \sum_{t'=0}^{k-1} (\|\hat{\theta}^{t'}\| + \sqrt{S^t} + 1) = \epsilon B \sum_{t'=0}^{k-1} (y^{t'} + 1)
\]
\[
\leq \epsilon B \sum_{t'=0}^{k-1} (2y^0 + 4\epsilon_k N B \tau_k + 1)
\]
\[
\leq \epsilon B \tau_k \left(2\|\bar{\theta}^0\| + 2\sqrt{S^0} + 2\right) \leq \frac{1}{3N}\left(\|\bar{\theta}^0\| + \sqrt{S^0} + 1\right)
\]
where again the second inequality comes from (35) and the last inequality follows from the condition on the step size.

\( \square \)

C. Proof of Lemma 3

We have shown in (38) that for all \( k \geq 0 \)
\[
\Theta^k - 1(\hat{\theta}^k)^T = (W - \frac{1}{N}11^T) \left((\Theta^{k-1} - 1(\hat{\theta}^{k-1})^T\right)
\]
\[
+ \epsilon_{k-1} \left(I - \frac{1}{N}11^T\right) F(X^{k-1}, \Theta^{k-1})
\]
As \( ||W - \frac{1}{N}11^T||_F \leq \sigma_2 \) and \( ||I - \frac{1}{N}11^T||_F \leq 1 \), we have
\[
S^k = ||\Theta^k - 1(\hat{\theta}^k)^T||_F^2
\]
\[
= \left\|W - \frac{1}{N}11^T\right\|_F^2 + \epsilon_{k-1} \left(I - \frac{1}{N}11^T\right) F(X^{k-1}, \Theta^{k-1})^2 \leq \left(1 + \frac{1 - \sigma_2^2}{4\sigma_2^2}\right) \|W - \frac{1}{N}11^T\|_F^2 + \epsilon_{k-1} \left(I - \frac{1}{N}11^T\right) F(X^{k-1}, \Theta^{k-1})^2 \leq \sigma_2^2 \left(1 + \frac{1 - \sigma_2^2}{4\sigma_2^2}\right) S^{k-1} + \epsilon_{k-1} \left(I - \frac{1}{N}11^T\right) F(X^{k-1}, \Theta^{k-1})^2 \leq \left(1 + \frac{1 - \sigma_2^2}{4\sigma_2^2}\right) \|W - \frac{1}{N}11^T\|_F^2 + \epsilon_{k-1} \left(I - \frac{1}{N}11^T\right) F(X^{k-1}, \Theta^{k-1})^2 \leq \frac{1 + 3\sigma_2^2}{4} S^{k-1} + \frac{8\epsilon_{k-1}^2}{1 - \sigma_2^2} \left(1 + 7\sigma_2^2\right) \sum_{i=1}^{N} ||F_i(X_i^{k-1}, \theta_i^{k-1})||^2 \leq \frac{1 + 3\sigma_2^2}{4} S^{k-1} + \frac{8\epsilon_{k-1}^2}{1 - \sigma_2^2} \sum_{i=1}^{N} B^2(\|\theta_i^{k-1}\| + 1)^2 \]
where in the first inequality, we use the fact that \( ||a + b||^2 \leq (1 + \frac{1}{\eta}) ||a||^2 + (1 + 2\eta) ||b||^2 \) for all vectors \( a, b \) and \( \eta > 0 \). The third

D. Proof of Lemma 4

To show Lemma 4, we introduce the following technical lemmas that bounds the cross term under Markovian data.

Lemma 5: We have for all \( k \geq \tau_k \)
\[
E[||\hat{\theta}^k - \theta^*||, \sum_{i=1}^{N} (F_i(X_i^k, \theta_i^k) - F_i(\bar{\theta}))]
\]
\[
\leq \frac{\alpha}{2} E[R^k] + M_1 \epsilon_k - \tau_k E[R_k] + M_2 E[S^k]
\]
\[
+ M_3 E[S^{k-\tau_k}] + \epsilon_{k-\tau_k} M_1 (||\theta^*||^2 + 1)
\]
where \( M_1 = \frac{45B + 30B^2 + 48BL}{N} \), \( M_2 = \frac{5B}{2} + \frac{NL^2}{2} + \frac{B}{2} \), \( M_3 = 5B + \frac{5NL^2}{N} + 4NL^2 + 5L \).
Proof of (Lemma 4): From the update rule (2)
\[
R^{k+1} \leq ||\hat{\theta}^k + \frac{\epsilon_k}{N} \sum_{i=1}^{N} F_i(X_i^k, \theta_i^k) - \theta^*||^2
\]
\[
= R^k + \frac{\epsilon_k}{N} \sum_{i=1}^{N} ||F_i(X_i^k, \theta_i^k)||^2 + 2\epsilon_k \langle \bar{\theta}^k - \theta^*, \sum_{i=1}^{N} F_i(\theta_i^k) \rangle
\]
\[
+ 2\epsilon_k \langle \bar{\theta}^k - \theta^*, \sum_{i=1}^{N} (F_i(X_i^k, \theta_i^k) - F_i(\bar{\theta})) \rangle
\]
\[
= R^k + \frac{\epsilon_k}{N} \sum_{i=1}^{N} ||F_i(X_i^k, \theta_i^k)||^2
\]
where the first equality follows from \( \sum_{i=1}^{N} \bar{F}_i(\theta^*) = 0 \), and the third inequality is a result of Assumption 3.

We bound the second term of (43) with Lemma 1

\[
\left\| \frac{\epsilon_k}{N} \sum_{i=1}^{N} F_i(X_i^k, \theta_i^k) \right\|^2 \\
\leq \frac{c_2^2}{N} \sum_{i=1}^{N} \left\| F_i(X_i^k, \theta_i^k) \right\|^2 \\
\leq \frac{c_2^2}{N} \sum_{i=1}^{N} B^2 (\| \theta_i^k \| + 1)^2 \\
\leq \frac{c_2^2 B^2}{N} \sum_{i=1}^{N} (3\| \theta_i^k \|^2 + 3\| \theta_i^k - \bar{\theta}_k \|^2 + 3) \\
\leq 3c_2^4 B^2 (\| \bar{\theta}_k \|^2 + 1) + \frac{3c_2^2 B^2}{N} S^k \\
\leq 6c_2^4 B^2 (R^k + \| \theta^* \|^2 + 1) + \frac{3c_2^2 B^2}{N} S^k. \tag{44}
\]

Lemma 5 bounds the third term of (43). This, along with (44) and (43), implies that

\[
\mathbb{E}[R^{k+1}] \\
\leq \mathbb{E} \left[ (1 - 2\alpha \epsilon_k) R^k + \left\| \frac{\epsilon_k}{N} \sum_{i=1}^{N} F_i(X_i^k, \theta_i^k) \right\|^2 + 2\epsilon_k \langle \bar{\theta}_k - \theta^* \rangle \sum_{i=1}^{N} (F_i(X_i^k, \theta_i^k) - \bar{F}_i(\bar{\theta}_k)) \right] \\
\leq (1 - 2\alpha \epsilon_k) \mathbb{E}[R^k] + 6c_2^2 B^2 \mathbb{E}[R^k] + \| \theta^* \|^2 + 1 \\
+ \frac{3c_2^2 B^2}{N} \mathbb{E}[S^k] + \alpha \epsilon_k \mathbb{E}[R^k] + 2M_1 \epsilon_k \epsilon_k - \tau_k \mathbb{E}[R^k] \\
+ 2M_2 \epsilon_k \mathbb{E}[S^k] + 2M_3 \epsilon_k \mathbb{E}[S^{k-\tau_k}] \\
+ 2\epsilon_k \epsilon_k - \tau_k M_1 (\| \theta^* \|^2 + 1) \\
\leq (1 - \alpha \epsilon_k) \mathbb{E}[R^k] + (2M_1 + 6B^2) \epsilon_k \epsilon_k - \tau_k \mathbb{E}[R^k] \\
+ \left( \frac{B}{2N^2} + 2M_2 \right) \epsilon_k \mathbb{E}[S^k] + 2M_3 \epsilon_k \mathbb{E}[S^{k-\tau_k}] \\
+ N C_1 \epsilon_k \epsilon_k - \tau_k \mathbb{E}[R^k] \\
\leq \left( 1 - \frac{\alpha \epsilon_k}{2} \right) \mathbb{E}[R^k] + N C_1 \epsilon_k \epsilon_k - \tau_k \mathbb{E}[R^k] \\
+ N \left( \frac{21B}{2} + 2L^2 \frac{\alpha}{\alpha} + L \right) \epsilon_k \mathbb{E}[S^k] \\
+ N \left( 10B + \frac{5}{6} + \frac{8L^2}{\alpha} + 10L \right) \epsilon_k \mathbb{E}[S^{k-\tau_k}] \\
\leq \left( 1 - \frac{\alpha \epsilon_k}{2} \right) \mathbb{E}[R^k] + N C_1 \epsilon_k \epsilon_k - \tau_k \mathbb{E}[R^k] \\
+ N C_2 \epsilon_k \mathbb{E}[S^k] + N C_2 \epsilon_k \mathbb{E}[S^{k-\tau_k}] \\
\right)
\]

where the third inequality simply follows from rearranging/grouping of the terms. In the fourth inequality, we use the condition that \( \epsilon_k - \tau_k \leq \frac{\alpha}{45N + 192N BL^2 + 192N BL} \), and \( C_1 \) and \( C_2 \) are defined in (9).

\[ \square \]

E. Proof of Lemma 5

Define

\[
T_{1}^k = \left\langle \bar{\theta}_k - \bar{\theta}^{k-\tau_k}, \sum_{i=1}^{N} (F_i(X_i^k, \theta_i^k) - \bar{F}_i(\bar{\theta}_k)) \right\rangle \\
T_{2}^k = \left\langle \bar{\theta}^{k-\tau_k} - \theta^*, \sum_{i=1}^{N} (F_i(X_i^k, \theta_i^{k-\tau_k}) - \bar{F}_i(\theta_i^{k-\tau_k})) \right\rangle \\
T_{3}^k = \left\langle \bar{\theta}_k - \theta^*, \sum_{i=1}^{N} (F_i(X_i^k, \theta_i^k) - F_i(X_i^k, \theta_i^{k-\tau_k})) \right\rangle \\
T_{4}^k = \left\langle \theta^{k-\tau_k} - \theta^*, \sum_{i=1}^{N} (\bar{F}_i(\theta_i^{k-\tau_k}) - \bar{F}_i(\bar{\theta}_k)) \right\rangle \\
T_{5}^k = \left\langle \bar{\theta}_k - \theta^*, \sum_{i=1}^{N} (\bar{F}_i(\bar{\theta}_k) - \bar{F}_i(\bar{\theta}_k)) \right\rangle. \tag{45}
\]

We introduce the following two lemmas to assist the proof of Lemma 5 and note that they are essentially consequences of the Lipschitz continuity of \( F_i \) and \( \bar{F}_i \).

Lemma 6: We have for all \( k \geq \tau_k \)

\[
T_{1}^k \leq \frac{30\epsilon_k - \tau_k B^2}{N} (N R^k + N S^{k-\tau_k} + S^k + N (\| \theta^* \|^2 + 1)) \tag{46}
\]

Lemma 7: We have for all \( k \geq \tau_k \)

\[
T_{5}^k + T_{4}^k + T_{5}^k \leq \frac{48 \epsilon_k - \tau_k N BL R^k + \alpha R^k}{2} \\
+ \frac{45 \epsilon_k - \tau_k N BL}{2} \left( 2 \| \theta^* \|^2 + 1 \right) \\
+ \left( \frac{NL^2}{\alpha} + \frac{L^2}{\alpha} \right) S^k \tag{47}
\]

Proof (of Lemma 5): To simplify notations, we define the norm of the local consensus error \( r^k_i = \| \theta_i^k - \bar{\theta}_k \| \). Note that \( r^k_i \)
relates to $S^k$
\[
\sum_{i=1}^{N} (r^k_i)^2 = S^k.
\]

From the definition of $S^k$ and $R^k$, we have $\sqrt{S^k} = \|\Theta^k - \mathbf{1}(\hat{\theta}^k)^\top\|_F$ and $\sqrt{R^k} = \|\theta^k - \theta^*\|$. 

We can decompose the quantity of interest using (45)
\[
\left\langle \hat{\theta}^k - \theta^*, \sum_{i=1}^{N} \left( F_i(X_i^k, \theta_i^k) - \bar{F}_i(\hat{\theta}^k) \right) \right\rangle = \mathcal{T}_1^k + \mathcal{T}_2^k + \mathcal{T}_3^k + \mathcal{T}_4^k.
\]

We define $\mathcal{F}_k = \{ \mathbf{X}_0, \ldots, \mathbf{X}_k, \Theta^0, \ldots, \Theta^k \}$. Recall the definition of the TV distance: for any two probability measures $a, b$ on a set $E$
\[
d_{TV}(a, b) = \frac{1}{2} \sup_{f \in E} \left| \int f \, da - \int f \, db \right|.
\]

This implies that
\[
\mathbb{E} \left[ \mathcal{T}_2^k \mid \mathcal{F}_k - \tau_k \right] = \left\langle \hat{\theta}^k - \tau_k - \theta^*, \sum_{i=1}^{N} \left( F_i(X_i^k, \theta_i^{k-\tau_k}) - \bar{F}_i(\theta_i^{k-\tau_k}) \right) \right\rangle.
\]

Using the geometric mixing time and $k \geq \tau_k$, we have $d_{TV}(P(X_i^k = .), \mu_i) \leq \epsilon_k$. This yields
\[
\mathbb{E} \left[ \mathcal{T}_2^k \mid \mathcal{F}_k - \tau_k \right] \leq B_{\epsilon_k} \sum_{i=1}^{N} \left( \|\hat{\theta}_{k-\tau_k} - \hat{\theta}^k\| + \sqrt{R^k} \right)
\times \left( r_{k-\tau_k}^k + \|\hat{\theta}^k - \hat{\theta}_{k-\tau_k}\| + \|\hat{\theta}^k\| + 1 \right)
\leq B_{\epsilon_k} \sum_{i=1}^{N} \left( \frac{1}{2N} \|\hat{\theta}^k\| + \frac{1}{2N} \sqrt{S^k-\tau_k} + \frac{1}{2N} + \sqrt{R^k} \right)
\times \left( r_{k-\tau_k}^k + \frac{3}{2} \|\hat{\theta}^k\| + \frac{1}{2N} \sqrt{S^k-\tau_k} + \frac{3}{2} \right)
\leq B_{\epsilon_k} \sum_{i=1}^{N} \left( 3\sqrt{R^k} + \frac{1}{N} \sqrt{S^k-\tau_k} + \frac{1}{N} + \frac{\|\theta^*\|}{N} \right)
\times \left( 2r_{k-\tau_k}^k + 3\sqrt{R^k} + 3\|\theta^*\| + \frac{1}{N} \sqrt{S^k-\tau_k} + 3 \right)
\leq B_{\epsilon_k} \sum_{i=1}^{N} \left( 3\sqrt{R^k} + 3\|\theta^*\| + 2r_{k-\tau_k}^k + \frac{1}{N} \sqrt{S^k-\tau_k} + 3 \right)^2
\]

where the second inequality is from (20) of Lemma 2, and the first and the third inequalities use the triangular inequality. 

The rest of the terms in (45) are bounded by Lemmas 6 and 7. Putting the terms together
\[
\mathbb{E} \left[ \left\langle \hat{\theta}^k - \theta^*, \sum_{i=1}^{N} \left( F_i(X_i^k, \theta_i^k) - \bar{F}_i(\hat{\theta}^k) \right) \right\rangle \right] \leq \mathbb{E} \left[ \mathcal{T}_1^k + \mathcal{T}_2^k + \mathcal{T}_3^k + \mathcal{T}_4^k \right]
\leq 30\epsilon_{k-\tau_k}^k \epsilon_k N B^2 \left( N \mathbb{E}[R^k] + N S^k - \tau_k + S^k + N\|\theta^*\|^2 + N \right) + \frac{5B_{\epsilon_k}}{4} N \left( 9R^k + 9\|\theta^*\|^2 + \frac{1}{N} S^k - \tau_k + 9 \right)
\]

where the third inequality simply regroups the terms, and the last inequality follows from $\epsilon_{k-\tau_k}^k \leq \frac{1}{2N\epsilon_k}$. 

\[\square\]
\[ T_1^k \leq \sum_{i=1}^{N} \| \bar{\theta}^k - \bar{\theta}^{k-\tau_k} \| \| F_i(X^k_i, \theta^k_i) - F_i(\bar{\theta}^k) \| \]
\[ \leq B \sum_{i=1}^{N} \| \bar{\theta}^k - \bar{\theta}^{k-\tau_k} \| (\| \theta^k_i \| + \| \bar{\theta}^k \| + 2) \]
\[ \leq 3\epsilon_{k-\tau_k} B^2 (\| \bar{\theta}^k \| + \sqrt{S^{k-\tau_k}} + 1) \sum_{i=1}^{N} (r^k_i + 2\| \bar{\theta}^k \| + 2) \]
\[ \leq 6\epsilon_{k-\tau_k} B^2 \sum_{i=1}^{N} (\sqrt{R^k} + \| \theta^k_i \| + \sqrt{S^{k-\tau_k}} + r^k_i + 1)^2 \]
\[ \leq 30\epsilon_{k-\tau_k} B^2 \sum_{i=1}^{N} (R^k + S^{k-\tau_k} + (r^k_i)^2 + \| \theta^k_i \|^2 + 1) \]
\[ = 30\epsilon_{k-\tau_k} B^2 (NR^k + NS^{k-\tau_k} + S^k + N(\| \theta^k \|^2 + 1)) \]

where the second inequality results from Lemma 1, and the third and fifth inequalities use the triangular inequality. The third inequality also uses (20) of Lemma 2. \( \square \)

\section*{G. Proof of Lemma 7}

Due to the Lipschitz continuity in Assumption 2 and the Cauchy–Schwarz inequality
\[ T_3^k + T_4^k + T_5^k \]
\[ \leq \sqrt{R^{k-\tau_k}} \sum_{i=1}^{N} L \left( r^k_i + r^{k-\tau_k} + \| \bar{\theta}^{k-\tau_k} - \bar{\theta}^k \| \right) v \]
\[ \leq L \left( \| \bar{\theta}^{k-\tau_k} - \bar{\theta}^k \| + \sqrt{R^k} \right) \]
\[ \times \sum_{i=1}^{N} \left( r^k_i + 2r^{k-\tau_k} + 2\| \bar{\theta}^{k-\tau_k} - \bar{\theta}^k \| \right) \]
\[ \leq L \sum_{i=1}^{N} \left( 3\epsilon_{k-\tau_k} B \| \bar{\theta}^k \| + 3\epsilon_{k-\tau_k} B \sqrt{S^{k-\tau_k}} + 3\epsilon_{k-\tau_k} B \| \bar{\theta}^k \| \right) \]
\[ + 3\epsilon_{k-\tau_k} B \sqrt{S^{k-\tau_k}} + 6\epsilon_{k-\tau_k} B \| \bar{\theta}^k \| \]
\[ \times \left( r^k_i + 2r^{k-\tau_k} + 6\epsilon_{k-\tau_k} B \| \bar{\theta}^k \| \right) \]
\[ + 6\epsilon_{k-\tau_k} B \sqrt{S^{k-\tau_k}} + 6\epsilon_{k-\tau_k} B \| \bar{\theta}^k \| \]
\[ = 3\epsilon_{k-\tau_k} B L \sum_{i=1}^{N} \left( \| \bar{\theta}^k \| + \sqrt{S^{k-\tau_k}} + 1 \right) (r^k_i + 2r^{k-\tau_k}) \]
\[ + 2L(3\epsilon_{k-\tau_k} B)^2 \sum_{i=1}^{N} \left( \| \bar{\theta}^k \| + \sqrt{S^{k-\tau_k}} + 1 \right)^2 \]

where the second inequality uses the triangular inequality, and the third inequality follows from (20) of Lemma 2. The last inequality simply regroups the terms. Then, using the Cauchy–Schwarz inequality and the fact that \( a \cdot b \leq \frac{a^2}{2} + \frac{b^2}{2} \) for any scalars \( a, b, c > 0 \), we can further simplify
\[ T_3^k + T_4^k + T_5^k \leq 3\epsilon_{k-\tau_k} B L \frac{\| \bar{\theta}^k \|^2 + S^{k-\tau_k} + 1}{2} \]
\[ + 6\epsilon_{k-\tau_k} B L \sum_{i=1}^{N} \| \bar{\theta}^k - \theta^* \| \left( \| \bar{\theta}^k \| + \sqrt{S^{k-\tau_k} + 1} \right) \]
\[ + 6\epsilon_{k-\tau_k} B L \sum_{i=1}^{N} \| \bar{\theta}^k - \theta^* \| \left( \| \bar{\theta}^k \| + \sqrt{S^{k-\tau_k} + 1} \right) \]
\[ + 6\epsilon_{k-\tau_k} B L \sum_{i=1}^{N} \| \bar{\theta}^k - \theta^* \| \left( \| \bar{\theta}^k \| + \sqrt{S^{k-\tau_k} + 1} \right) \]

where the last two inequality simply regroup the terms and simplify the expressions using the conditions on the step size. \( \square \)

\section*{References}

[1] H. Robbins and S. Monro, “A stochastic approximation method,” Ann. Math. Statist., vol. 22, no. 3, pp. 400–407, 1951.

[2] A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro, “Robust stochastic approximation approach to stochastic programming,” SIAM J. Optim., vol. 19, no. 4, pp. 1574–1609, 2009.

[3] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
[4] K. Ovchinnikov, A. Semakova, and A. Matveev, “Decentralized multi-agent tracking of unknown environmental level sets by a team of non-holonomic robots,” in Proc. 6th Int. Congr. Ultra Modern Telecommun. Control Syst., 2014, pp. 352–359.

[5] T. Zeng, T. T. Doan, and J. Romberg, “Finite-time analysis of decentralized stochastic approximation with applications in multi-agent and multi-task learning,” in Proc. IEEE Conf. Decis. Control, 2021, pp. 2641–2646.

[6] X. Lian, C. Zhang, H. Zhang, C.-J. Hsieh, W. Zhang, and J. Liu, “Can decentralized algorithms outperform centralized algorithms? A case study for decentralized parallel stochastic gradient descent;” in Proc. Adv. Neural Inf. Process. Syst., 2017, pp. 5330–5340.

[7] J. Zeng and W. Yin, “On nonconvex decentralized gradient descent,” IEEE Trans. Signal Process., vol. 66, no. 11, pp. 2834–2848, Jun. 2018.

[8] A. Koloskova, N. Loizou, S. Boreiri, M. Jaggi, and S. Stich, “A unified theory of decentralized SGD with changing topology and local updates,” in Proc. Int. Conf. Mach. Learn., 2020, pp. 5381–5390.

[9] A. Khaled, K. Mischchenko, and P. Richtarik, “Tightst theory for local SGD on identical and heterogeneous data,” in Proc. Int. Conf. Artif. Intell. Statist., 2020, pp. 4519–4529.

[10] T. Sun, T. Chen, Y. Sun, Q. Liao, and D. Li, “Decentralized Markov chain gradient descent,” 2019, arXiv:1909.10238.

[11] H.-T. Wai, “On the convergence of consensus algorithms with Markovian noise and gradient bias,” in Proc. IEEE Conf. Decis. Control, 2020, pp. 4897–4902.

[12] A. Khaled and P. Richtarik, “Better theory for SGD in the nonconvex world,” 2020, arXiv:2002.03329.

[13] S. K. Rikriat and L. Ying, “Finite-time error bounds for linear stochastic approximation and TD learning,” in Proc. Conf. Learn. Theory, 2019, pp. 2803–2830.

[14] Z. Chen, S. Zhang, T. T. Doan, S. T. Maguluri, and J.-P. Clarke, “Finite-sample analysis of nonlinear stochastic approximation with applications in reinforcement learning,” 2019, arXiv:1905.11425.

[15] T. Doan, S. Maguluri, and J. Romberg, “Finite-time analysis of distributed TD (0) with linear function approximation on multi-agent reinforcement learning,” in Proc. Int. Conf. Mach. Learn., 2019, pp. 1626–1635.

[16] T. T. Doan, S. T. Maguluri, and J. Romberg, “Finite-time performance of distributed temporal-difference learning with linear function approximation,” SIAM J. Math. Data Sci., vol. 3, no. 1, pp. 298–320, 2021.

[17] K. Zhang, Z. Yang, H. Liu, T. Zhang, and T. Basar, “Fully decentralized multi-agent reinforcement learning with networked agents,” in Proc. Int. Conf. Mach. Learn., 2018, pp. 5872–5881.

[18] M. Assran, J. Romoff, B. J. Pinaeu, and M. Rabbat, “Gossip-based actor-learner architectures for deep reinforcement learning,” in Proc. Adv. Neural Inf. Process. Syst., 2019, pp. 13320–13330.

[19] S. Zeng, M. A. Anwar, T. T. Doan, A. Raychowdhury, and J. Romberg, “A decentralized policy gradient approach to multi-task reinforcement learning,” in Proc. Uncertainty Artif. Intell., 2021, pp. 1002–1012.

[20] Y. Lin, G. Qi, L. Huang, and A. Wierman, “Multi-agent reinforcement learning in stochastic networked systems,” in Proc. Adv. Neural Inf. Process. Syst., 2021, pp. 7825–7837.

[21] S. Zeng, T. Chen, A. Garcia, and M. Hong, “Learning to coordinate in multi-agent systems: A coordinated actor-critic algorithm and finite-time guarantees,” in Proc. Learn. Dyn. Control Conf., 2022, pp. 278–290.

[22] K. Zhang, Z. Yang, H. Liu, T. Zhang, and T. Baatar, “Finite-sample analysis for decentralized batch multiaejgnt reinforcement learning with networked agents,” IEEE Trans. Autom. Control, vol. 66, no. 12, pp. 5925–5940, Dec. 2021.

[23] P. Kairouz et al., “Advances and open problems in federated learning,” Found. Trends Mach. Learn., vol. 14, no. 1/2, pp. 1–210, 2021.

[24] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated learning: Challenges, methods, and future directions,” IEEE Signal Process. Mag., vol. 37, no. 3, pp. 50–60, May 2020.

[25] T. T. Doan, “Local stochastic approximation: A unified view of federated learning and distributed multi-task reinforcement learning algorithms,” 2020, arXiv:2006.13460.

[26] D. A. Levin, Y. Peres, and E. L. Wilmer, Markov Chains and Mixing Times. American Math. Soc., 2006.

[27] S. Ram, A. Nedić, and V. V. Veeravalli, “Incremental stochastic subgradient algorithms for convex optimization,” SIAM J. Optim., vol. 20, no. 2, pp. 691–717, 2009.

[28] H. Karimi, J. Nutini, and M. Schmidt, “Linear convergence of gradient and proximal-gradient methods under the Polyak-Lojasiewicz condition,” in Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discov. Databases, 2016, pp. 795–811.

[29] A. Nedić, A. Olshevsky, and M. G. Rabbat, “Network topology and communication-computation tradeoffs in decentralized optimization,” Proc. IEEE, vol. 106, no. 5, pp. 953–976, May 2018.

[30] M. O. Sayin, N. D. Vanli, S. S. Koza, and T. Baatar, “Stochastic subgradient algorithms for strongly convex optimization over distributed networks,” IEEE Trans. Netw. Sci. Eng., vol. 4, no. 4, pp. 248–260, Apr. 2017.

[31] J. Hale and S. Lunel, Introduction to Functional Differential Equations, vol. 99, Berlin, Germany: Springer, 1993.

[32] A. Nedić and A. Olshevsky, “Distributed optimization over time-varying directed graphs,” IEEE Trans. Autom. Control, vol. 60, no. 3, pp. 601–615, Mar. 2015.

[33] T. Sun, Y. Sun, and W. Yin, “On Markov chain gradient descent,” in Proc. Adv. Neural Inf. Process. Syst., 2018, pp. 9896–9905.

[34] B. Poljak and T. Tskykin, “Robust identification,” Automatica, vol. 16, no. 1, pp. 53–63, 1980.

[35] J. Kim and J. P. Lynch, “Autonomous decentralized system identification by Markov parameter estimation using distributed smart wireless sensor networks,” J. Eng. Mechanics, vol. 135, no. 5, pp. 478–490, 2010.

[36] O. Ige, “Markov chain epidemic models and parameter estimation,” Ph.D. dissertation, Dept. Math., Marshall Univ., Huntington, WV, USA, 2020.

[37] Y. Zhou, L. Wang, R. Zhong, and Y. Tan, “A Markov chain based demand prediction model for stations in bike sharing systems,” Math. Problems Eng., vol. 2018, 2018.

[38] A. Olshevsky, “Linear time average consensus on fixed graphs,” IFAC-PapersOnLine, vol. 48, no. 22, pp. 94–99, 2015.

Sihan Zeng (Graduate Student Member, IEEE) received the B.S. degree in electrical and computer engineering and the B.A. degree in statistics from Rice University, Houston, TX, USA, in 2017. He is currently working toward the Ph.D. degree in machine learning with the Georgia Institute of Technology, Atlanta, GA, USA. His current research interest includes optimization, reinforcement learning, and applied probability.

Thinh T. Doan (Member, IEEE) received the bachelor and master's degree from the Hanoi University of Science and Technology, Hanoi, Vietnam, the master degree from the University of Oklahoma, Norman, OK, USA, and the Ph.D. degree from the University of Illinois, Urbana-Champaign, Champaign, IL, USA, in 2008, 2013, and 2018, all in electrical engineering. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA. His research interests include the intersection of control theory, optimization, machine learning, and applied probability.

Justin Romberg (Fellow, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical and computer engineering from Rice University, Houston, TX, USA, in 1997, 1999, and 2004, respectively.

He is currently the Schlumberger Professor with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA, where he has been on the faculty since 2006. He was a Postdoctoral Scholar in applied and computational mathematics with the California Institute of Technology. His current research interests include the intersection of statistical signal processing, machine learning, optimization, and applied probability.