A LIST OF CONSTANTS

In this appendix, we list all the constants used in our main results, Theorems 3 and 5. They are finite and their expressions do not affect the understanding of the theorems. Since their expressions are quite long and complicated, we begin with the following set of constants, based on which we will be able to present the constants used in the theorems and the proofs of the theorems in an easier way. We hope that this way can also help the readers to better understand and follow our results and analyses.

The first constant ζ_1 is defined as follows. Recall that ϵ is given in equation 6 as

$$\epsilon = \left(1 + \frac{2b_{\max}}{A_{\max}} - \frac{\pi_{\min}\beta^{2L}}{2\delta_{\max}}\right)(1+\alpha A_{\max})^{2L} - \frac{2b_{\max}}{A_{\max}}(1+\alpha A_{\max})^{L}.$$

ζ_1 is defined as the unique solution for which $\epsilon = 1$ if $\alpha = \zeta_1$. The following remark shows why ζ_1 uniquely exists.

Remark 5 *From equation 6, it is easy to see that ϵ is monotonically increasing for $\alpha > 0$. Define the corresponding monotonic function as*

$$f(\alpha) = \left(1 + \frac{2b_{\max}}{A_{\max}} - \frac{\pi_{\min}\beta^{2L}}{2\delta_{\max}}\right)(1+\alpha A_{\max})^{2L} - \frac{2b_{\max}}{A_{\max}}(1+\alpha A_{\max})^{L}.$$

Note that $0 < f(0) < 1$ and $f(+\infty) = +\infty$. Thus, $f(\alpha) = 1$ has a unique solution ζ_1. $\square$

The other constants are defined as follows:

$$\zeta_2 = \frac{4b_{\max}^2}{A_{\max}^2}\left[(1+\alpha A_{\max})^L - 1\right]^2 + 2b_{\max}\frac{(1+\alpha A_{\max})^L - 1}{A_{\max}}(1+\alpha A_{\max})^L \tag{12}$$

$$\begin{aligned}\zeta_3 = &\left(144 + 4A_{\max}^2 + 912\tau(\alpha)A_{\max}^2 + 168\tau(\alpha)A_{\max}b_{\max}\right)\|\theta^*\|_2^2 \\ &+ \tau(\alpha)A_{\max}^2\left[152\left(\frac{b_{\max}}{A_{\max}} + \|\theta^*\|_2\right)^2 + \frac{48b_{\max}}{A_{\max}}\left(\frac{b_{\max}}{A_{\max}} + 1\right)^2 + \frac{87b_{\max}^2}{A_{\max}^2} + \frac{12b_{\max}}{A_{\max}}\right] \\ &+ 2 + 2b_{\max}^2 + 4\|\theta^*\|_2^2 + \frac{48b_{\max}^2}{A_{\max}^2}\end{aligned} \tag{13}$$

$$\zeta_4 = \sqrt{N}b_{\max}\left(2 + \frac{12b_{\max}^2}{A_{\max}^2} + 38\|\theta^*\|_2^2\right) \tag{14}$$

$$\zeta_5 = 144 + 916A_{\max}^2 + 168A_{\max}b_{\max} \tag{15}$$

$$\zeta_6 = 4b_{\max}^2\alpha L^2(1+\alpha A_{\max})^{2L-2} + 2b_{\max}L(1+\alpha A_{\max})^{2L-1} \tag{16}$$

$$\begin{aligned}\zeta_7 = &(148 + 916A_{\max}^2 + 168A_{\max}b_{\max})\|\theta^*\|_2^2 + 2 + \frac{48b_{\max}^2}{A_{\max}^2} + 152\left(b_{\max} + A_{\max}\|\theta^*\|_2\right)^2 \\ &+ 89b_{\max}^2 + 12A_{\max}b_{\max} + 48A_{\max}b_{\max}\left(\frac{b_{\max}}{A_{\max}} + 1\right)^2\end{aligned} \tag{17}$$

$$\zeta_8 = 144 + 916A_{\max}^2 + 168A_{\max}b_{\max} + 144A_{\max}\mu_{\max} \tag{18}$$

$$\begin{aligned}\zeta_9 = &\Bigg[2 + (4+\zeta_8)\|\theta^*\|_2^2 + 48\frac{(b_{\max}+\mu_{\max})^2}{A_{\max}^2} + 152\left(b_{\max} + \mu_{\max} + A_{\max}\|\theta^*\|_2\right)^2 \\ &+ 12A_{\max}b_{\max} + 48A_{\max}(b_{\max}+\mu_{\max})\left(\frac{b_{\max}+\mu_{\max}}{A_{\max}} + 1\right)^2 + 89(b_{\max}+\mu_{\max})^2\Bigg]\end{aligned} \tag{19}$$

Here $\mu_{\max} = (N+1)A_{\max}C_\theta$, where C_θ is a finite number defined in Lemma 19 which can be regarded as an upper bound of 2-norm of each agent i's state θ_t^i generated by the Push-SA algorithm equation 9.

A.1 CONSTANTS USED IN THEOREM 3

$$\begin{aligned}
K_1 &= \min\left\{\zeta_1,\ \frac{\gamma_{\max}}{0.9}\right\} \\
K_2 &= 144 + 4A_{\max}^2 + 912\tau(\alpha)A_{\max}^2 + 168\tau(\alpha)A_{\max}b_{\max} \\
C_1 &= \frac{\gamma_{\max}}{\gamma_{\min}}\left(8\exp\left\{2\alpha A_{\max}T_1\right\}+4\right)\mathbf{E}\left[\|\langle\theta\rangle_0-\theta^*\|_2^2\right] \\
&\quad + 8\frac{\gamma_{\max}}{\gamma_{\min}}\exp\left\{2\alpha A_{\max}T_1\right\}\left(\|\theta^*\|_2+\frac{b_{\max}}{A_{\max}}\right)^2 \\
C_2 &= \frac{2\zeta_2}{1-\epsilon}+\frac{\gamma_{\max}}{\gamma_{\min}}\cdot\frac{2\alpha\zeta_3\gamma_{\max}}{0.9} \\
C_3 &= \frac{2\zeta_6}{1-\epsilon} \\
C_4 &= 2\zeta_7\alpha_0 C\frac{\gamma_{\max}}{\gamma_{\min}} \\
C_5 &= 2\alpha_0\zeta_4\frac{\gamma_{\max}}{\gamma_{\min}} \\
C_6 &= 2LT_2\frac{\gamma_{\max}}{\gamma_{\min}}\mathbf{E}\left[\|\langle\theta\rangle_{LT_2}-\theta^*\|_2^2\right]
\end{aligned} \tag{20}$$

T_1 is any positive integer such that for all $t \geq T_1$, there hold $t \geq \tau(\alpha)$ and $36\sqrt{N}b_{\max}\eta_{t+1}\gamma_{\max} + K_2\alpha\gamma_{\max} \leq 0.1$.

Remark 6 *We show that T_1 must exist. From $0 < \alpha < \min\{K_1,\ \frac{\log 2}{A_{\max}\tau(\alpha)},\ \frac{0.1}{K_2\gamma_{\max}}\}$, it is easy to see that the feasible set of α is nonempty and $K_2\alpha\gamma_{\max} < 0.1$. Since $\lim_{t\to\infty}\eta_t = 0$ by Lemma 9 and $\tau(\alpha) \leq -C\log\alpha$ by Assumption 3, there exists a time instant $T \geq -C\log\alpha$ such that for any $t \geq T$, there hold $t \geq \tau(\alpha)$ and $\eta_{t+1} \leq (0.1 - K_2\alpha\gamma_{\max})/(36\sqrt{N}b_{\max}\gamma_{\max})$, which implies that T_1 exists.* □

T_2 is any positive integer such that for all $t \geq LT_2$, there hold $\alpha_t \leq \alpha$, $2\tau(\alpha_t) \leq t$, $\tau(\alpha_t)\alpha_{t-\tau(\alpha_t)} \leq \min\{\frac{\log 2}{A_{\max}},\ \frac{0.1}{\zeta_5\gamma_{\max}}\}$ and $\zeta_5\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max} + 36\sqrt{N}b_{\max}\eta_{t+1}\gamma_{\max} \leq 0.1$.

Remark 7 *We explain why T_2 must exist. Since $\alpha_t = \frac{\alpha_0}{t+1}$ is monotonically decreasing for t and $\tau(\alpha_t) \leq -C\log\alpha_t = -C\log\alpha_0 + C\log(t+1)$ from Assumption 3, there exists a positive S_1 such that for any $t \geq S_1$, we have $\alpha_t \leq \alpha$ and $t \geq 2\tau(\alpha_t)$ for any constant $0 < \alpha < \min\{K_1,\ \frac{\log 2}{A_{\max}\tau(\alpha)},\ \frac{0.1}{K_2\gamma_{\max}}\}$. Moreover, it is easy to show that*

$$\begin{aligned}
\lim_{t\to\infty} t - \tau(\alpha_t) &\geq \lim_{t\to\infty} t + C\log\alpha_0 - C\log(t+1) = +\infty, \\
\lim_{t\to\infty} \tau(\alpha_t)\alpha_{t-\tau(\alpha_t)} &\leq \lim_{t\to\infty} \frac{-C\alpha_0\log\alpha_0 + C\alpha_0\log(t+1)}{t-\tau(\alpha_t)+1} = 0.
\end{aligned}$$

Then, there exists a positive S_2 such that for any $t \geq S_2$, we have $\tau(\alpha_t)\alpha_{t-\tau(\alpha_t)} \leq \min\{\frac{\log 2}{A_{\max}},\ \frac{0.1}{\zeta_5\gamma_{\max}}\}$. In addition, since $\lim_{t\to\infty}\eta_t = 0$ from Lemma 9, when $\tau(\alpha_t)\alpha_{t-\tau(\alpha_t)} \leq \frac{0.1}{\zeta_5\gamma_{\max}}$, there exists a positive S_3 such that for any $t \geq S_3$, we have $\eta_{t+1} \leq (0.1 - \zeta_5\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max})/(36\sqrt{N}b_{\max}\gamma_{\max})$. Therefore, T_2 must exist as we can simply set $T_2 = \max\{S_1,\ S_2,\ S_3\}$. □

A.2 CONSTANTS USED IN THEOREM 5

$$
\begin{aligned}
C_7 &= \frac{16}{\epsilon_1}\mathbf{E}\left[\Big\|\sum_{i=1}^{N}\tilde{\theta}_0^i + \alpha_0 A(X_0)\tilde{\theta}_0^i + \alpha_0 b^i(X_0)\Big\|_2\right] \\
C_8 &= \frac{16}{\epsilon_1}\cdot\frac{A_{\max}C_\theta + b_{\max}}{1-\bar{\epsilon}} \\
C_9 &= 2A_{\max}C_\theta + 2b_{\max} \\
C_{10} &= 2N\zeta_9\alpha_0 C\frac{\gamma_{\max}}{\gamma_{\min}} \\
C_{11} &= 2\alpha_0 N\frac{\gamma_{\max}}{\gamma_{\min}} \\
C_{12} &= 2\bar{T}N\frac{\gamma_{\max}}{\gamma_{\min}}\mathbf{E}\left[\|\langle\tilde{\theta}\rangle_{\bar{T}} - \theta^*\|_2^2\right]
\end{aligned}
$$

Here ϵ_1 is a positive constant defined as $\epsilon_1 = \inf_{t\geq 0}\min_{i\in\mathcal{V}}(\hat{W}_t\cdots\hat{W}_0\mathbf{1}_N)^i$. From Corollary 2 (b) in Nedić & Olshevsky (2015) and the fact that each $\hat{W}_t$ is column stochastic, $\epsilon_1 \in [\frac{1}{N^{NL}}, 1]$. See Lemma 20 for more details.

$\bar{T}$ is any positive integer such that for all $t \geq \bar{T}$, there hold $2\tau(\alpha_t) \leq t$, $\mu_t + \tau(\alpha_t)\alpha_{t-\tau(\alpha_t)}\zeta_8 \leq \frac{0.1}{\gamma_{\max}}$ and $\tau(\alpha_t)\alpha_{t-\tau(\alpha_t)} \leq \min\{\frac{\log 2}{A_{\max}}, \frac{0.1}{\zeta_8\gamma_{\max}}\}$.

Remark 8 *From Lemma 21,* $\lim_{t\to\infty}\mu_t = 0$*. Then, using the similar arguments as in Remark 7, we can show the existence of* $\bar{T}$*.* □

B RELATED WORK

A key tool used for designing and analyzing RL algorithms is stochastic approximation (Robbins & Monro, 1951), e.g., for policy evaluation, including temporal difference (TD) learning as a special case (Sutton & Barto, 2018). Convergence study of stochastic approximation based on ordinary differential equation (ODE) methods has a long history (Borkar & Meyn, 2000). Notable examples are Tsitsiklis & Roy (1997); Dayan (1992) which prove asymptotic convergence of TD(λ). Recently, *finite-time performance* of single-agent stochastic approximation and TD algorithms has been studied in Dalal et al. (2018a); Lakshminarayanan & Szepesvari (2018); Bhandari et al. (2018); Srikant & Ying (2019); Gupta et al. (2019); Wang et al. (2017); Ma et al. (2020); Xu et al. (2019); Chen et al. (2020b); many other works have now appeared that perform finite-time analysis for other RL algorithms, see, e.g., Zou et al. (2019); Qu & Wierman (2020); Wu et al. (2020); Xu & Gu (2020); Weng et al. (2020); Wang & Zou (2020); Chen et al. (2020a); Wang et al. (2019); Dalal et al. (2018b); Borkar & Pattathil (2018), just to name a few.

C DISCUSSION ON ASSUMPTION 6

In this appendix, we contend that Assumption 6 has more general applications than the previously known case and that it is in fact necessary.

C.1 APPLICATIONS

First, as mentioned in Remark 3, there are at least two cases which satisfy Assumption 6, yet cannot be directly handled by the existing analysis tool, which was developed only for doubly stochastic matrices. Case 1 is when the number of in-neighbors of agents is unchanged over time. This case has an interesting behavioral interpretation in fish biology, and has been adopted in bio-inspired distributed algorithm design (Abaid & Porfiri, 2010). Case 2 is when the interaction matrix changes arbitrarily over time during an initial period, after which it finally becomes fixed. As we describe below, Case 2 occurs naturally in certain multi-agent systems.

Case 1 is mathematically equivalent to the situation when all stochastic matrices share the same left dominant eigenvector, which subsumes doubly stochastic matrices as a special case; thus it could be analyzed by carefully choosing a fixed norm. There may be different choices: one choice is to apply our time-varying quadratic Lyapunov comparison function $\sum_{i=1}^{N} \pi_t^i \mathbf{E}[||\theta_t^i - \theta^*||_2^2]$ to the time-invariant case (i.e., π_t^i does not change over time), which leads to the weighted Frobenius norm defined in the appendix.

The extension to Case 1 just described may be straightforward, but Case 2 is not. As we proved in Theorems 2 and 3, when the interaction matrix arbitrarily changes over time for an initial period, say of length T, and finally becomes a fixed matrix or enters Case 1, all agents' trajectories determined by (1) will converge in mean square. Also, recall that the corresponding finite-time error bounds in this case were derived using the "absolute probability sequence" technique. Note that the existing techniques can only be applied to analyze (1) after time T; when T is very large, such an analysis is undesirable, since the focus and challenge here are for "finite" time.

It is important to note that Case 2 provides a realistic model for certain systems. Consider scenarios in which some agents do not function stably and thus they communicate with their neighbors sporadically for a certain period, leading to a time-varying stochastic matrix. Such scenarios occur naturally when there is unstable communication due to environmental changes or movement of agents (e.g., robots or UAVs may need to move into a new formation while continuing computation). After this unstable period, which could be long, the whole system then enters a stable operation status. This satisfies Case 2 and our finite-time analysis can be applied to the whole process, no matter how long the unstable period could be, as long as it is finite. In addition to this example, Case 2 and our analysis can be applied to certain scenarios in the presence of malicious agents. Suppose the system is aware that a small subset of agents have potentially been attacked and are thus behaving maliciously. To protect the system, the consensus interaction among the agents can switch to resilient consensus algorithms such as Vaidya et al. (2012); LeBlanc et al. (2013) in order to attenuate the effect of malicious agents. In this situation, the resulting dynamics of the non-malicious agents are in general characterized by a time-varying stochastic matrix. After identifying and/or fixing the malicious agents, which could be a very slow process, the system can switch back to normal operation status. This example again satisfies Case 2, and our analysis can be applied to the whole procedure. As we mentioned in Remark 3, if some malicious agents always exist, the non-malicious agents in general will not converge, and thus a finite-time analysis is probably meaningless. The non-convergence issue will be further explained in the next subsection.

Whether Assumption 6 can represent more realistic/analytic examples is a very interesting future direction. Though consensus has been extensively studied and the "absolute probability sequence" was proposed decades ago, this question has never been explored. The development of more advanced analysis tools is an interesting topic as well.

C.2 NECESSITY

We now elaborate on why Assumption 6 is not restrictive from a theoretical point of view.

As mentioned in Remark 3, distributed SA with time-varying stochastic matrices does not converge, in general, if Assumption 6 does not hold. Assumption 6 is sufficient to guarantee the convergence of the distributed SA algorithm equation 1 when the interaction matrix is row stochastic and time-varying. Let us denote the necessary and sufficient condition for convergence of consensus-based distributed SA as Condition A, which is currently unknown. It is possible that there is a large gap between Assumption 6 and Condition A. But Assumption 6 is (to our knowledge) the most general sufficient condition that has been proposed so far; one indirect justification of this claim is Assumption 6 is an analogue of condition (C3.4') in Kushner & Yin (1987), which is itself a sufficient condition guaranteeing the asymptotic convergence of a different form of distributed SA. While Kushner & Yin (1987) only provided asymptotic analysis, we provided both asymptotic and finite-time analyses using a novel tool. Assumption 6 subsumes the existing analysis for doubly stochastic matrices as a special case, and can be used for more general, nontrivial cases (see the examples provided in the discussion of Case 2 above). Existing analysis tools cannot be applied to Case 2. From a theoretical point of view, our paper reduces the gap between the doubly stochastic matrices assumption and Condition A to the smaller gap between Assumption 6 and Condition A, for finite-time analysis of consensus-based distributed SA.

In addition, the other equally important main contribution of our paper, push-SA, does not need Assumption 6, though its analysis still relies on the "absolute probability sequence" technique.

C.3 CONTRIBUTIONS

Next, we present a high-level view of our paper, which may help the readers to better understand our overall contributions.

There are three major information fusion schemes in the vast distributed algorithms literature: "consensus" (time-varying stochastic matrices), "averaging" (time-varying doubly stochastic matrices which include gossiping), and "push-sum" (time-varying column stochastic matrices). The consensus-based scheme can guarantee an agreement among the agents, but the agreement point in general cannot be specified, especially when the interaction is time-varying. The averaging scheme can specify the agreement point to be the average among all agents using doubly stochastic matrices, but these only work for undirected graphs (i.e., bi-directional communication is required between any pair of neighbors); typical examples are the Metropolis algorithm (Xiao et al., 2005) and gossiping (Boyd et al., 2006). The push-sum scheme is able to not only achieve agreement on the average, but it also works for directed graphs, allowing uni-directional communication. The push-sum scheme can also be straightforwardly modified to achieve any given convex combination agreement among all agents. The three schemes are widely used, depending on task specifications. Push-sum appears to be the most powerful, but the other two also have advantages; e.g., consensus can be modified to be more resilient against malicious agents, and averaging is easier in algorithm design (especially gossiping) and analysis (due to nicer properties of doubly stochastic matrices). There is a very recently proposed scheme called push-pull, but it is not yet that popular, so we focus our attention on the three major schemes.

With the above background in mind, there are three major information fusion schemes that can be used to design distributed SA (as well as RL). The existing literature has only analyzed the averaging scheme (doubly stochastic matrices), which to us appears to be the easiest among the three. Finite-time analyses of the other two schemes are untouched in the literature. Our paper is the first to consider both.

As explained in the preceding subsection, our result and analysis for the consensus scheme (based on Assumption 6) are the most general so far and generalize the existing tools in a nontrivial manner. This leads to very interesting, open research problems – like necessary and sufficient condition for distributed SA convergence – as well as how to design resilient consensus fusion, which can guarantee convergence of distributed SA.

D DISTRIBUTED TD LEARNING

In this section, we apply our distributed stochastic approximation finite-time analyses to distributed TD learning, as TD(λ) is a special cases of stochastic approximation. To this end, we first introduce the following multi-agent MDP tailored for distributed TD.

The multi-agent MDP can be defined by a tuple $(\mathcal{S}, \{\mathcal{U}^i\}_{i\in\mathcal{V}}, \{R^i\}_{i\in\mathcal{V}}, \bar{P}, \gamma, \{\mathbb{G}_t\}_{t\geq 0})$, where $\mathcal{S} = \{1, \dots, S\}$ is the finite set of S states, $\mathcal{U}^i$ is the set of control actions for agent i. For each agent i, $R^i : \mathcal{S}\times\mathcal{U}\times\mathcal{S} \to \mathbb{R}$ is the local reward function, where $\mathcal{U} = \prod_{i=1}^N \mathcal{U}^i$ is the joint control action space. $\bar{P} : \mathcal{S}\times\mathcal{U}\times\mathcal{S} \to [0, 1]$ denotes the state transition probability matrix of the MDP, and $\gamma \in (0, 1)$ is the discount factor. Given a fixed policy, we let $\bar{P}$ be of size $S \times S$ for convenience, and thus its ij-th entry $\bar{p}^{ij}$ equals the probability from state i to state j under the given policy. The multi-agent MDP then evolves as follows. At each time $t \geq 0$, each agent i observes the current state $s_t \in \mathcal{S}$, takes action $u_t^i = \mu^i(s_t) \in \mathcal{U}^i$, and receive a corresponding reward $R^i(s_t, u_t, s_{t+1})$, where $\mu^i : \mathcal{S} \to \mathcal{U}^i$ is a function mapping a state to a control action in $\mathcal{U}^i$ and $u_t = \prod_{i=1}^N u^i \in \mathcal{U}$. It is worth emphasizing that in such a multi-agent setting, each agent's rewards and reward function are private information, and thus cannot be shared with any other agents.

The discounted accumulative reward $J : \mathcal{S} \to \mathbb{R}$ associated with the multi-agent MDP is defined for each $s \in \mathcal{S}$ as

$$J(s) = \mathbf{E}\left[\sum_{t=0}^{\infty} \gamma^t \sum_{i\in\mathcal{V}} c^i R^i(s_t, u_t, s_{t+1}) \mid s_0 = s\right], \tag{21}$$

which satisfies the Bellman equation Sutton & Barto (2018), i.e.,

$$J(s) = \sum_{s'=1}^{S} \bar{p}^{ss'} \left[\sum_{i\in\mathcal{V}} c^i R^i(s, s') + \gamma J(s')\right], \quad s \in \mathcal{S},$$

where $c^i > 0$, $i \in \mathcal{V}$, is a set of convex combination weights. The existing distributed RL algorithms all set $c^i = 1/N$ for all $i \in \mathcal{V}$, e.g. Zhang et al. (2018); Doan et al. (2019), and this is why they require interaction matrices all be doubly stochastic. We will show that $c^i = \pi_\infty^i$ for all $i \in \mathcal{V}$ for general stochastic matrix sequences. Since for any doubly stochastic matrix sequence, its absolute probability sequence is $\pi_t = (1/N)\mathbf{1}_N$, i.e., $\pi_\infty^i = 1/N$ for all $i \in \mathcal{V}$, our results generalize the existing results, e.g. Doan et al. (2019; 2021). In § 3, we will show how to achieve the straight average reward, i.e.,$c^i = 1/N$ for all $i \in \mathcal{V}$, without requiring doubly stochastic matrices.

When the number of the states is very large, the computation of exact J may be intractable. To get around this, as did in Tsitsiklis & Roy (1997), we use a low-dimensional linear function $\hat{J}$ to approximate J. Specifically, the linear function approximator $\hat{J}$ takes the form

$$\hat{J}(s, \theta) = \sum_{k=1}^{K} \theta^k \phi_k^s, \quad s \in \mathcal{S},$$

where each ϕ_k^s is a fixed scalar function defined on the state space $\mathcal{S}$, each θ^k is the associated weight, and $K \ll S$. In other words, $\hat{J}$ is parameterized by $\theta \in \mathbb{R}^K$, with θ^k being the k-th entry of θ. To proceed, let $\phi_k \in \mathbb{R}^S$ be the vector whose j-th entry is ϕ_k^j for all $k \in \{1, \dots, K\}$, let $\phi(s) \in \mathbb{R}^K$ be the vector whose j-th entry is ϕ_j^s for all $s \in \mathcal{S}$, and let $\Phi \in \mathbb{R}^{S\times K}$ be the matrix whose i-th row is the row vector $\phi(i)^\top$ and whose j-th column is the vector ϕ_j, i.e.,

$$\Phi = \begin{bmatrix} \phi_1 & \cdots & \phi_K \end{bmatrix} = \begin{bmatrix} \phi(1)^\top \\ \vdots \\ \phi(S)^\top \end{bmatrix} \in \mathbb{R}^{S\times K},$$

which implies $\hat{J} = \Phi\theta$. The goal for the multi-agent network is to find an optimal θ^* with which the distance between $\hat{J}$ and J is minimized, under the following standard assumptions adopted e.g. Srikant & Ying (2019); Doan et al. (2019).

Assumption 7 *All the rewards are uniformly bounded, i.e., there exists a positive constant R such that $|R^i(s, s')| \leq R$ for all $i \in \mathcal{V}$ and $s, s' \in \mathcal{S}$.*

Assumption 8 *The vectors $\phi_1, \dots, \phi_K$ are linearly independent, i.e., Φ has full column rank, and $\|\phi(s)\|_2 \leq 1$ for all $s \in \mathcal{S}$.*

Assumption 9 *The Markov chain that evolves according to the transition probability matrix $\bar{P}$ is irreducible and aperiodic. Let $d \in \mathbb{R}^S$ be the unique stationary distribution associated with $\bar{P}$, i.e., $d^\top \bar{P} = d^\top$.*

D.1 DISTRIBUTED TD(λ)

In this subsection, we make use of TD(λ) to estimate θ^* in a distributed manner. Note that, TD(0) can be applied in the similar manner. Each agent i updates its own estimator of θ^*, θ_t^i, as follows:

$$\theta_{t+1}^i = \sum_{j\in\mathcal{N}_t^i} w_t^{ij} \theta_t^j + \alpha_t \left(A(X_t) \sum_{j\in\mathcal{N}_t^i} w_t^{ij} \theta_t^j + b^i(X_t)\right), \quad i \in \mathcal{V}, \quad t \in \{0, 1, 2, \dots\}, \tag{22}$$

where $X_t = (s_t, s_{t+1}, z_t)$ is the Markov chain, with $z_t = \sum_{k=0}^{t} (\gamma\lambda)^{t-k}\phi(s_k)$, and

$$A(X_t) = z_t(\gamma\phi(s_{t+1}) - \phi(s_t))^\top, \quad b^i(X_t) = r_t^i z_t, \tag{23}$$

with r_t^i being the reward for agent i at time t. It is worth emphasizing that the proposed TD(λ) algorithm is different from that in Doan et al. (2021). The update equation 22 with equation 23 is a special case of equation 1.

In the sequel, we will show that the update equation 22 with equation 23 is a special case of equation 1 so that our analysis for equation 1 can be applied here. To this end, let $D = \text{diag}(d) \in \mathbb{R}^{S\times S}$, where d is given in Assumption 9,

$$A = \Phi^\top D(U - I)\Phi, \quad U = (1-\lambda)\sum_{t=0}^{\infty} \lambda^t(\gamma \bar{P})^{t+1}, \quad b^i = \Phi^\top D \sum_{t=0}^{\infty} (\gamma\lambda\bar{P})^t r^i, \quad i \in \mathcal{V}, \tag{24}$$

where $r^i \in \mathbb{R}^S$ whose k-th entry is $r^{ik} = \sum_{s=1}^{S} \bar{p}^{ks} R^i(k,s)$, and set $A_{\max} = \frac{1+\gamma}{1-\gamma\lambda}$ and $b_{\max} = \frac{R}{1-\gamma\lambda}$, where R is given in Assumption 7.

Lemma 2 *Let the sequences $\{\theta_t^i\}$, $i \in \mathcal{V}$, be generated by equation 22 with equation 23. If Assumptions 7–9 hold, so do Assumptions 2–4.*

Proof of Lemma 2: Firstly, under Assumptions 7–9, we have

$$\lim_{t\to\infty} \mathbf{E}[A(X_t)] = A, \quad \lim_{t\to\infty} \mathbf{E}[B(X_t)] = \begin{bmatrix} (b^1)^\top \\ \vdots \\ (b^N)^\top \end{bmatrix},$$

and $\|A(X_t)\|_2 \le \frac{1+\gamma}{1-\gamma\lambda}$, $\quad \|b(X_t)\|_2 \le \frac{R}{1-\gamma\lambda}$, where $A(X_t)$ and $b^i(X_t)$ are defined in equation 23, A and b^i are defined in equation 24. Since $A_{\max} = 1+\gamma$ and $b_{\max} = R$, then we know that Assumption 2 has been satisfied. Moreover,

$$\begin{aligned}
&\|\mathbf{E}[b^i(X_t) - b^i|S_0 = s_0, S_1 = s_1]\|_2 \\
&= \|\sum_{s=1}^{S} (\mathbf{P}(S_t = s|S_1 = s_1) - \pi_s)\phi(s) r^{is}\|_2 \le b_{\max}\sum_{i=s}^{S} |\mathbf{P}(S_t = i|S_1 = s_1) - \pi_i|, \\
&\|\mathbf{E}[A(X_t) - A|S_0 = s_0, S_1 = s_1]\|_2 \\
&= \|\sum_{s=1}^{S} (\mathbf{P}(S_t = s|S_1 = s_1) - \pi_s)\phi(s)(\sum_{j=1}^{S} \bar{p}^{sj}\gamma\phi(j)^\top - \phi(s)^\top)\|_2 \\
&\le A_{\max}\sum_{i=1}^{S} |\mathbf{P}(S_t = s|S_1 = s_1) - \pi_i|.
\end{aligned}$$

Since $\{S_t\}$ is a finite state, aperiodic and irreducible Markov chain, it has a geometric mixing rate (Brémaud, 2013), which implies that Assumption 3 holds. Lastly, when Assumption 8 holds, from the proof of Theorem 1 in Tsitsiklis & Roy (1997), A given in equation 24 is a negative definite matrix, i.e., $x^\top Ax < 0$ for all $x \in \mathbb{R}^K$, which implies that $A + A^\top$ is a symmetric negative definite matrix. From Theorem 7.11 in Rugh (1996), A is a Hurwitz matrix. ■

Then, using the similar arguments as in Lemma 2, we can show that Assumptions 7–9 imply Assumptions 2–4, and thus our analysis for equation 1 can be applied here. From the proof of Theorem 1 in Tsitsiklis & Roy (1997), A in equation 24 is a negative definite matrix, which implies that $A + A^\top$ is a symmetric negative definite matrix. Thus, we can also choose $P = I$ in Assumption 4 and use the Lyapunov function $V(\langle\theta\rangle_t) = \|\langle\theta\rangle_t - \theta^*\|_2^2$ in the analysis, where θ^* here is the limiting point of equation 22. Using the same argument as in Theorem 2, we can show that θ^* is the unique equilibrium point of the ODE equation 5 with A and b^i being defined in equation 24.

The finite-time performance of the distributed TD(λ) algorithm is characterized by the following theorem.

Theorem 6 *Let the sequences $\{\theta_t^i\}$, $i \in \mathcal{V}$, be generated by equation 22 with equation 23. Suppose that Assumptions 1 and 6–9 hold and $\{\mathbb{G}_t\}$ is uniformly strongly connected by sub-sequences of length L. Let δ_t be the diameter of $\cup_{k=t}^{t+L-1}\mathbb{G}_k$ and $\delta = \max_{t\geq 0}\delta_t$. Set $A_{\max} = \frac{1+\gamma}{1-\gamma\lambda}$, $b_{\max} = \frac{R}{1-\gamma\lambda}$, and $\sigma_{\min} > 0$ be the smallest eigenvalue of $-\frac{1}{2}(A + A^\top)$, where A is given in equation 24. Let $0 < \alpha < \min\left\{\Psi_9, \frac{\log 2}{A_{\max}\tau(\alpha)}, \frac{\sigma_{\min}}{\Psi_3}, \frac{1}{\sigma_{\min}}\right\}$.*

1) Fixed step-size: *Let $\alpha_t = \alpha$ for all $t \geq 0$. For all $t \geq \hat{T}_1$,*

$$\sum_{i=1}^N \pi_t^i \mathbf{E}\left[\|\theta_t^i - \theta^*\|_2^2\right] \leq 2\epsilon^{q_t}\sum_{i=1}^N \pi_{m_t}^i \mathbf{E}\left[\|\theta_{m_t}^i - \langle\theta\rangle_{m_t}\|_2^2\right] + (1-\alpha\sigma_{\min})^{t-\hat{T}_1}\hat{C}_1 + \hat{C}_2 + \hat{C}_3\sum_{k=0}^{t-\hat{T}_1}\eta_{t+1-k}(1-\alpha\sigma_{\min})^k. \tag{25}$$

2) Time-varying step-size: *Let $\alpha_t = \frac{\alpha_0}{t+1}$ with $\alpha_0 \geq \frac{1}{\sigma_{\min}}$. For all $t \geq \hat{T}_1 L$,*

$$\sum_{i=1}^N \pi_t^i \mathbf{E}[\|\theta_t^i - \theta^*\|_2^2] \leq 2\epsilon^{q_t-\hat{T}_1}\sum_{i=1}^N \pi_{\hat{T}_1 L+m_t}^i \mathbf{E}\left[\|\theta_{\hat{T}_1 L+m_t}^i - \langle\theta\rangle_{\hat{T}_1 L+m_t}\|_2^2\right] + \hat{C}_4\left(\alpha_0\epsilon^{\frac{q_t-1}{2}} + \alpha_{\lceil\frac{q_t-1}{2}\rceil L}\right) + \frac{1}{t}\left(\hat{C}_5\log^2\left(\frac{t}{\alpha_0}\right) + \hat{C}_6\sum_{l=\hat{T}_1 L}^t \eta_l + \hat{C}_7\right). \tag{26}$$

Here $\hat{T}_1, \hat{T}_1 - \hat{C}_7$ are finite constants whose definitions are given in Appendix E.2 with $A_{\max} = \frac{1+\gamma}{1-\gamma\lambda}$ and $b_{\max} = \frac{R}{1-\gamma\lambda}$.

D.2 PUSH-TD(λ)

In this subsection, we propose a push-based distributed TD(λ) algorithm and provide its finite-time error bounds. Note that, push-based distributed TD(0) can be applied in the similar manner. Each agent $i \in \mathcal{V}$ updates its variables at each time $t \geq 0$ as follows:

$$\begin{cases} y_{t+1}^i = \sum_{j\in\mathcal{N}_t^i}\hat{w}_t^{ij}y_t^j, \quad y_0^i = 1, \\ \hat{\theta}_{t+1}^i = \sum_{j\in\mathcal{N}_t^i}\hat{w}_t^{ij}\hat{\theta}_t^j + \alpha_t\Big(A(X_t)\hat{w}_t^{ij}\theta_t^j + b^j(X_t)\Big), \\ \theta_{t+1}^i = \dfrac{\hat{\theta}_{t+1}^i}{y_{t+1}^i}, \end{cases} \tag{27}$$

where $\hat{w}_t^{ij} = 1/|\mathcal{N}_t^{j-}|$, $X_t = (s_t, s_{t+1}, z_t)$ is the Markov chain, with $z_t = \sum_{k=0}^t(\gamma\lambda)^{t-k}\phi(s_k)$, $A(X_t)$ and $b^i(X_t)$ are given in equation 23. Using the same argument as in Theorem 4, we can show that θ^* is the unique equilibrium point of the ODE equation 10 with A and b^i being defined in equation 24.

Theorem 7 *Suppose that Assumptions 7–9 hold and $\{\mathbb{G}_t\}$ is uniformly strongly connected by sub-sequences of length L. Let the sequences $\{\theta_t^i\}$, $i \in \mathcal{V}$, be generated by equation 27 with equation 23, $\alpha_t = \frac{\alpha_0}{t+1}$ and $\alpha_0 \geq \frac{1}{\sigma_{\min}}$. Then, there exists a nonnegative $\bar{\epsilon} \leq (1-\frac{1}{N^{NL}})^{\frac{1}{L}}$ such that for all $t \geq \bar{T}$,*

$$\sum_{i=1}^N \mathbf{E}\left[\left\|\theta_{t+1}^i - \theta^*\right\|_2^2\right] \leq C_7\bar{\epsilon}^t + C_8\left(\alpha_0\bar{\epsilon}^{\frac{t}{2}} + \alpha_{\lceil\frac{t}{2}\rceil}\right) + C_9\alpha_t + \frac{1}{t}\left(C_{10}\log^2\left(\frac{t}{\alpha_0}\right) + C_{11}\sum_{k=\bar{T}}^t\mu_k + C_{12}\right). \tag{28}$$

Here $\bar{T}$ and $C_7 - C_{12}$ are finite constants whose definitions are given in Appendix in Appendix A with $A_{\max} = \frac{1+\gamma}{1-\gamma\lambda}$, $b_{\max} = \frac{R}{1-\gamma\lambda}$ and $\gamma_{\max} = \gamma_{\min} = 1$.

D.3 Constants for TD

$$\hat{C}_1 = \left(8\exp\left\{2\alpha A_{\max}\hat{T}_1\right\}+4\right)\mathbf{E}\left[\|\langle\theta\rangle_0-\theta^*\|_2^2\right]+8\exp\left\{2\alpha A_{\max}\hat{T}_1\right\}\left(\|\theta^*\|_2+\frac{b_{\max}}{A_{\max}}\right)^2;$$
$$\hat{C}_2 = \frac{2\zeta_2}{1-\epsilon}+\frac{2\alpha\zeta_3}{\sigma_{\min}};$$
$$\hat{C}_3 = 2\alpha\zeta_4;$$
$$\hat{C}_4 = \frac{2\zeta_6}{1-\epsilon};$$
$$\hat{C}_5 = 2\zeta_7\alpha_0 C;$$
$$\hat{C}_6 = 2\alpha_0\zeta_4;$$
$$\hat{C}_7 = 2\hat{T}_2 L\mathbf{E}\left[\|\langle\theta\rangle_{\hat{T}_2 L}-\theta^*\|_2^2\right].$$

$\hat{T}_1$ is any positive integer such that for all $t \geq \hat{T}_1$, there hold $t \geq \tau(\alpha)$ and $36\sqrt{N}b_{\max}\eta_{t+1}+K_2\alpha \leq \sigma_{\min}$.

$\hat{T}_2$ is any positive integer such that for all $t \geq \hat{T}_2 L$, there hold $\alpha_t \leq \alpha$, $2\tau(\alpha_t) \leq t$, $\tau(\alpha_t)\alpha_{t-\tau(\alpha_t)} \leq \min\{\frac{\log 2}{A_{\max}}, \frac{\sigma_{\min}}{\zeta_5}\}$ and $\zeta_5\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)+36\sqrt{N}b_{\max}\eta_{t+1} \leq \sigma_{\min}$.

D.4 Simulations

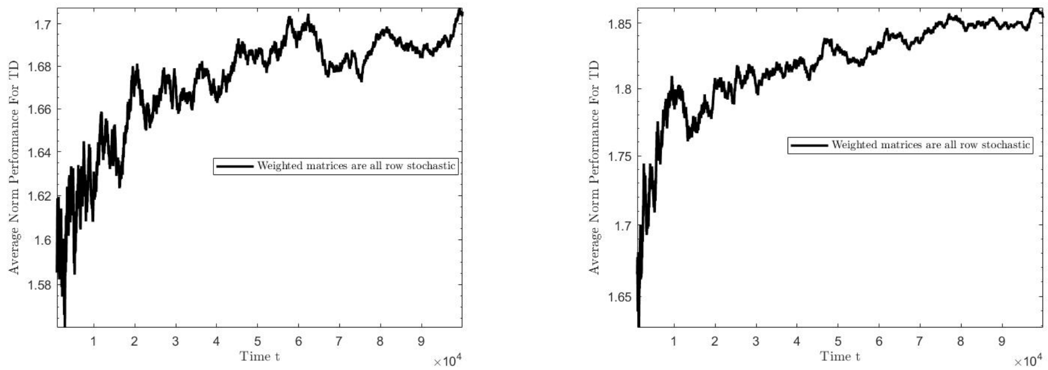
In this section, we numerically validate the finite-time bounds derived in this paper, for both distributed TD(λ) and push-TD(λ), and compare with the existing distributed TD(λ) results in Doan et al. (2021). We focus on TD(λ) and TD(0) as the existing distributed TD(0) finite-time analysis in Doan et al. (2019) only considers i.i.d. samples.

The TD setting and multi-agent network are given as follows. Set $\lambda = 0.2$ and discount factor $\gamma = 0.3$. Consider an environment with $N = 10$ agents and $|\mathcal{S}| = 10$ states. We generated a row stochastic matrix with each entry in $[0,1]$ and then added a small constant 10^{-5} to each element to make sure that the transition matrix satisfies Assumption 9. For each agent i and state pair (s,s'), we randomly sampled mean reward $R^i(s,s')$ from $[-3,3]$, and the instantaneous reward r_t^i was randomly sampled from $[R^i(s,a)-0.5, R^i(s,a)+0.5]$. The dimension of the feature vector ϕ was set as $K = 5$. We sampled the entry of ϕ from $[0,1]$ while simultaneously guaranteeing that the feature matrix Φ satisfies Assumption 8.

First, we considered consensus-based algorithm equation 22 with time-varying stochastic matrices to show the necessity of Assumption 6. To this end, we simulated two cases. In this first case, we randomly generated a stochastic matrix at each time step, and thus the corresponding absolute probability sequence π_t does not converge. We set the time-varying stepsize as $\alpha_t = 1/t^{0.68}$. Figure 1 (a) shows that in this case the average norm of all agents variables does not converge, implying non-convergence of all agents' states. In the second case, we consider a more special case in which the underlying graph changes periodically. In many distributed algorithms like distributed optimization, periodic settings can be regarded as a time-invariant case which thus guarantees convergence. We set the period as 10 and constructed the same set of 10 different stochastic matrices for each period. However, Figure 1 (b) shows that even with this periodic setting, the consensus-based algorithm (1) still does not converge, because a periodic sequence of stochastic matrices does not have a convergent absolute probability sequence.

Next, we will compare the finite-time bounds derived in this paper with the one in the existing literature Doan et al. (2021). As mentioned earlier, compared with the finite-time analysis of distributed TD(λ) with doubly stochastic matrices in Doan et al. (2021), the finite-time analysis in the paper is more general. In the sequel, we will evaluate the theoretical finite-time bounds for both distributed TD(λ) and push-TD(λ). To illustrate the differences and advantage over Doan et al. (2021), for consensus-based distributed TD(λ), we consider the following three settings:

1. The first 15 weight matrices are fixed, row stochastic (not doubly stochastic) and the following 85 weight matrices are fixed, doubly stochastic (see Figure 2 (a) - (d)).



(a) Random row stochastic matrices

(b) Periodic row stochastic matrices

Figure 1: Non-convergent distributed TD(λ) algorithm (equation 22) without Assumption 6.

Table 1: Comparison of the asymptotic bounds

	Bound in this paper	Bound in Doan et al. (2021)
Mixed stochastic matrices* + fixed stepsize	88.1653	114.1386
Mixed stochastic matrices + time-varying stepsize	0	0
All doubly stochastic matrices + fixed stepsize	881.5952	1141.4
All non-doubly, stochastic matrices + fixed stepsize	241.6035	NA
All column stochastic matrices (push-TD) + time-varying stepsize	0	NA

*A set of mixed stochastic matrices contains both doubly and non-doubly ones

2. All weight matrices are fixed, stochastic matrices, but not doubly stochastic (see Figure 2 (e), (f)).

3. All weight matrices are fixed, doubly stochastic (see Figure 2 (g), (h)).

In addition, we evaluated the finite-time bound of the push-TD(λ) algorithm with a fixed, column stochastic weight matrix (see Figure 3).

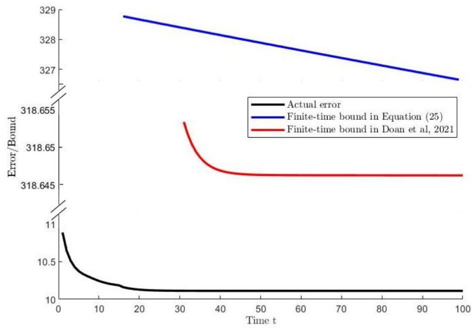
Figure 2 (a), 2 (c) and 2 (e) show that the bounds in equation 25, equation 26 and Doan et al. (2021) can be used to bound the actual error. However, the stating time for the bound in equation 25 and equation 26 are earlier than those in Doan et al. (2021).

Figure 2 (b) and 2 (f) show the both bounds in equation 25 and Doan et al. (2021) will converge to some fixed values when the stepsize is fixed. These values are listed and compared in Table 1. Figure 2(d) shows that both the bounds in equation 26 and Doan et al. (2021) will converge to zero for the time-varying stepsize.

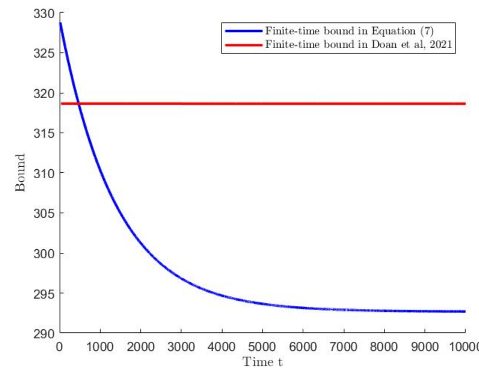
Figure 2 (g) and 2 (h) show that the bound in Doan et al. (2021) can not be (directly) applied to the stochastic (not doubly stochastic) weight matrix case. In addition, Figure 2 (h) shows that the bound in equation 25 will converge to some fixed value when the stepsize is fixed, which is given in Table 1.

Figure 3 shows that the push-TD(λ) algorithm will converge to the optimal point in the long run, and the bound in equation 28 can be used to bound the actual error. It is worth emphasizing that the bound in Doan et al. (2021) cannot be applied for this case.

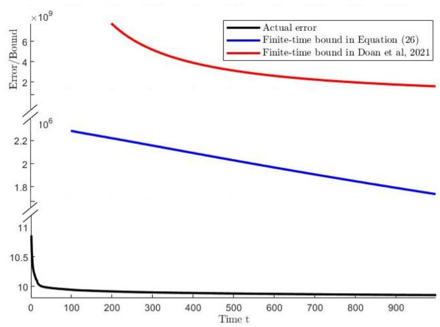
In summary, from the figures and Table 1, we can see that (1) our consensus-based TD(λ) can be applied to more general time-varying row stochastic matrices cases; though our finite-time bound is



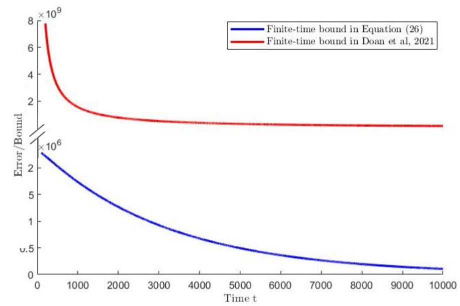
(a) Mixed stochastic matrices (fixed stepsize)



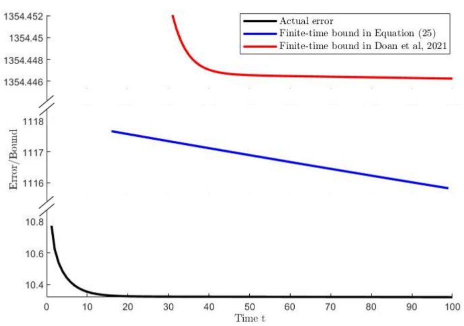
(b) Mixed stochastic matrices (fixed stepsize)



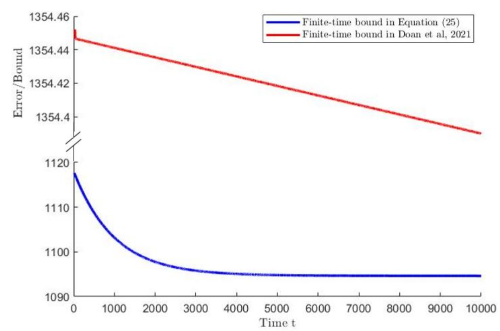
(c) Mixed stochastic matrices (time-varying stepsize)



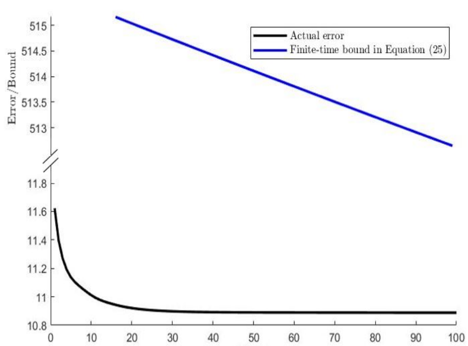
(d) Mixed stochastic matrices (time-varying stepsize)



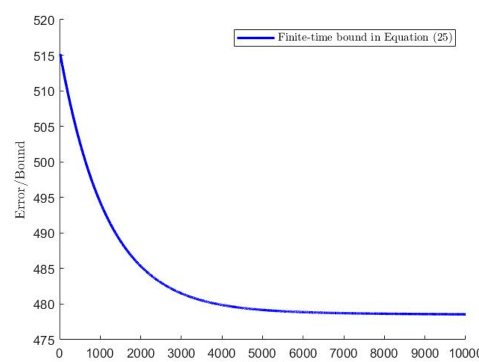
(e) All doubly stochastic matrices (fixed stepsize)



(f) All doubly stochastic matrices (fixed stepsize)



(g) All non-doubly, stochastic matrices (fixed stepsize)



(h) All non-doubly, stochastic matrices (fixed stepsize)

Figure 2: Finite-time bounds for consensus-based TD(λ) (equation 22).

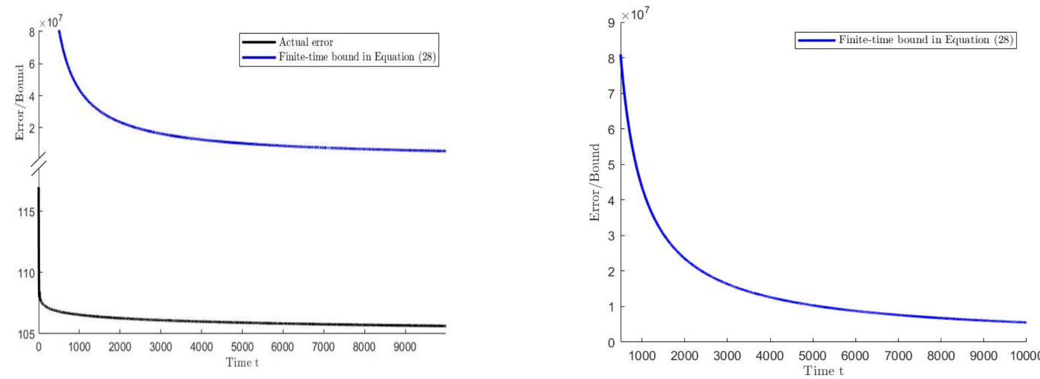


Figure 3: Finite-time bounds for push-TD(λ) (equation 27) with time-varying stepsizes.

looser at the beginning period, it can be applied at an earlier time instant and has a tighter limiting bound, compared with Doan et al. (2021); and (2) our push-based TD(λ) expands the applicability of the existing distributed TD learning, as it can work for any time-varying directed graphs as long as they are uniformly strongly connected, without any restrictive assumptions such as Assumption 6 in our consensus-based one and the doubly stochasticity assumption in Doan et al. (2021).

E ANALYSIS AND PROOFS

In this appendix, we provide the analysis of our two algorithms, equation 1 and equation 9, and the proofs of all the assertions in the paper. We begin with some notation.

E.1 NOTATION

We use $\mathbf{0}_n$ to denote the vector in $\mathbb{R}^n$ whose entries all equal to 0's. For any vector $x \in \mathbb{R}^n$, we use $\text{diag}(x)$ to denote the $n \times n$ diagonal matrix whose ith diagonal entry equals x^i. We use $\|\cdot\|_F$ to denote the Frobenius norm. For any positive diagonal matrix $W \in \mathbb{R}^{n\times n}$, we use $\|A\|_W$ to denote the weighted Frobenius norm for $A \in \mathbb{R}^{n\times m}$, defined as $\|A\|_W = \|W^{\frac{1}{2}}A\|_F$. It is easy to see that $\|\cdot\|_W$ is a matrix norm. We use $\mathbf{P}(\cdot)$ to denote the probability of an event and $\mathbf{E}(X)$ to denote the expected value of a random variable X.

E.2 DISTRIBUTED STOCHASTIC APPROXIMATION

In this subsection, we analyze the distributed stochastic approximation algorithm equation 1 and provide the proofs of the results in Section 2. We begin with the asymptotic performance.

Proof of Lemma 1: Since the uniformly strongly connectedness is equivalent to B-connectedness as discussed in Remark 2, the existence is proved in Lemma 5.8 of Touri (2012), and the uniqueness is proved in Lemma 1 of Nedić & Liu (2017). ■

Proof of Theorem 1: Without loss of generality, let $\{\mathbb{G}_t\}$ be uniformly strongly connected by sub-sequences of length L. Note that for any $i \in \mathcal{V}$, we have

$$0 \leq \pi_{\min}\|\theta_t^i - \langle\theta\rangle_t\|_2^2 \leq \pi_{\min}\sum_{j=1}^{N}\|\theta_t^j - \langle\theta\rangle_t\|_2^2 \leq \sum_{j=1}^{N}\pi_t^j\|\theta_t^j - \langle\theta\rangle_t\|_2^2, \tag{29}$$

where $\pi_{\min}$ is defined in Lemma 1. From Lemma 10,

$$
\begin{aligned}
&\lim_{t\to\infty}\sum_{i=1}^{N}\pi_t^i\|\theta_t^i-\langle\theta\rangle_t\|_2^2\\
&\le \lim_{t\to\infty}\hat{\epsilon}^{q_t-T_4^*}\sum_{i=1}^{N}\pi^i_{T_4^*L+m_t}\|\theta^i_{T_4^*L+m_t}-\langle\theta\rangle_{T_4^*L+m_t}\|_2^2+\lim_{t\to\infty}\frac{\zeta_6}{1-\hat{\epsilon}}\left(\alpha_0\hat{\epsilon}^{\frac{q_t-1}{2}}+\alpha_{\lceil\frac{q_t-1}{2}\rceil L}\right)\\
&=0. \qquad (30)
\end{aligned}
$$

Combining equation 29 and equation 30, it follows that for all $i\in\mathcal{V}$, $\lim_{t\to\infty}\pi_{\min}\|\theta_t^i-\langle\theta\rangle_t\|_2^2=0$. Since $\pi_{\min}>0$ by Lemma 1, $\lim_{t\to\infty}\|\theta_t^i-\langle\theta\rangle_t\|_2=0$ for all $i\in\mathcal{V}$. ■

Proof of Theorem 2: From Theorem 1, all θ_t^i, $i\in\mathcal{V}$, will reach a consensus with $\langle\theta\rangle_t$ and the update of $\langle\theta\rangle_t$ is given in equation 4, which can be treated as a single-agent linear stochastic approximation whose corresponding ODE is equation 5. From Kushner & Yin (1987); Kushner (1983),[1] we know that $\langle\theta\rangle_t$ will converge to θ^* w.p.1, which implies that θ_t^i will converge to θ^* w.p.1. In addition, from Theorem 3-(2) and Lemma 9, $\lim_{\to\infty}\sum_{i=1}^N\pi_t^i\mathbf{E}[\|\theta_t^i-\theta^*\|_2^2]=0$. Since π_t^i is uniformly bounded below by $\pi_{\min}>0$, as shown in Lemma 1, it follows that θ_t^i will converge to θ^* in mean square for all $i\in\mathcal{V}$. ■

We now analyze the finite-time performance of equation 1. In the sequel, we use K to denote the dimension of each θ_t^i, i.e., $\theta_t^i\in\mathbb{R}^K$ for all $i\in\mathcal{V}$.

E.2.1 Fixed Step-size

We first consider the fixed step-size case and begin with validation of two "convergence rates" in Theorem 3.

Lemma 3 *Both ϵ and $(1-\frac{0.9\alpha}{\gamma_{\max}})$ lie in the interval $(0,1)$.*

Proof of Lemma 3: Since $0<\alpha<K_1=\min\{\zeta_1,\ \frac{\gamma_{\max}}{0.9}\}$ as imposed in Theorem 3, we have $0<\alpha<\zeta_1$ and $0<\alpha<\frac{\gamma_{\max}}{0.9}$. The latter immediately implies that $1-\frac{0.9\alpha}{\gamma_{\max}}\in(0,1)$. From Remark 5, ϵ is monotonically increasing for $\alpha>0$. In addition, from the definition of ζ_1 in Section A that if $\alpha=\zeta_1$, then $\epsilon=1$. Since $0<\alpha<\zeta_1$, it follows that $0<\epsilon<1$. ■

To proceed, we need the following derivation and lemmas.

Let $Y_t=\Theta_t-\mathbf{1}_N\langle\theta\rangle_t^\top=(I-\mathbf{1}_N\pi_t^\top)\Theta_t$. For any $t\ge s\ge0$, let $W_{s:t}=W_tW_{t-1}\cdots W_s$. Then,

$$
\begin{aligned}
Y_{t+1}&=\Theta_{t+1}-\mathbf{1}_N\langle\theta\rangle_{t+1}^\top\\
&=W_t\Theta_t+\alpha W_t\Theta_tA^\top(X_t)+\alpha B(X_t)-\mathbf{1}_N(\langle\theta\rangle_t^\top+\alpha\langle\theta\rangle_t^\top A^\top(X_t)+\alpha\pi_{t+1}^\top B(X_t))\\
&=W_t(I-\mathbf{1}_N\pi_t^\top)\Theta_t+\alpha W_t(I-\mathbf{1}_N\pi_t^\top)\Theta_tA^\top(X_t)+\alpha(I-\mathbf{1}_N\pi_{t+1}^\top)B(X_t)\\
&=W_tY_t+\alpha W_tY_tA^\top(X_t)+\alpha(I-\mathbf{1}_N\pi_{t+1}^\top)B(X_t). \qquad (31)
\end{aligned}
$$

For simplicity, let Y_t^i be the i-th column of matrix $Y_t^\top$. Then,

$$
Y_{t+1}^i=\sum_{j=1}^N w_t^{ij}Y_t^j+\alpha A(X_t)\sum_{j=1}^N w_t^{ij}Y_t^j+\alpha\left(b^i(X_t)-B^\top(X_t)\pi_{t+1}\right). \qquad (32)
$$

[1] On page 1289 of Kushner & Yin (1987), it says that the idea in Kushner (1983) can be adapted to get the w.p.1 convergence result.

From equation 31, we have

$$\begin{aligned}
Y_{t+L} &= W_{t+L-1}Y_{t+L-1}(I+\alpha A^\top(X_{t+L-1}))+\alpha(I-\mathbf{1}_N\pi_{t+L}^\top)B(X_{t+L-1})\\
&= W_{t+L-1}W_{t+L-2}Y_{t+L-2}(I+\alpha A^\top(X_{t+L-2}))(I+\alpha A^\top(X_{t+L-1}))\\
&\quad+\alpha W_{t+L-1}(I-\mathbf{1}_N\pi_{t+L-1}^\top)B(X_{t+L-2})(I+\alpha A^\top(X_{t+L-1}))\\
&\quad+\alpha(I-\mathbf{1}_N\pi_{t+L}^\top)B(X_{t+L-1})\\
&= W_{t:t+L-1}Y_t(I+\alpha A^\top(X_t))\cdots(I+\alpha A^\top(X_{t+L-1}))+\alpha(I-\mathbf{1}_N\pi_{t+L}^\top)B(X_{t+L-1})\\
&\quad+\alpha\sum_{k=t}^{t+L-2}W_{k+1:t+L-1}(I-\mathbf{1}_N\pi_{k+1}^\top)B(X_k)\left(\Pi_{j=k+1}^{t+L-1}(I+\alpha A^\top(X_j))\right),
\end{aligned}\tag{33}$$

and

$$Y_{t+L}^i=\left(\Pi_{k=t}^{t+L-1}(I+\alpha A(X_k))\right)\sum_{j=1}^N w_{t:t+L-1}^{ij}Y_t^j+\alpha\hat{b}_{t+L}^i,$$

where

$$\begin{aligned}
\hat{b}_{t+L}^i &= (b^i(X_{t+L-1})-B(X_{t+L-1})^\top\pi_{t+L})\\
&\quad+\sum_{k=t}^{t+L-2}\left(\Pi_{j=k+1}^{t+L-1}(I+\alpha A(X_j))\right)\sum_{j=1}^N w_{k+1:t+L-1}^{ij}(b^j(X_k)-B(X_k)^\top\pi_{k+1}).
\end{aligned}$$

Lemma 4 *Suppose that Assumption 1 holds and $\{\mathbb{G}_t\}$ is uniformly strongly connected by sub-sequences of length L. Then, for all $t\geq 0$,*

$$\sum_{i=1}^N\pi_{t+L}^i\sum_{j=1}^N\sum_{k=1}^N w_{t:t+L-1}^{ij}w_{t:t+L-1}^{ik}\|Y_t^j-Y_t^k\|_2^2\geq\frac{\pi_{\min}\beta^{2L}}{\delta_{\max}}\sum_{i=1}^N\pi_t^i\|Y_t^i\|_2^2,$$

where $\beta>0$ and $\pi_{\min}>0$ are given in Assumption 1 and Lemma 1, respectively.

Proof of Lemma 4: We first consider the case when $K=1$, i.e., $Y_t^i\in\mathbb{R},\forall i$. From Lemma 1, we have

$$\begin{aligned}
&\sum_{i=1}^N\pi_{t+L}^i\sum_{j=1}^N\sum_{l=1}^N w_{t:t+L-1}^{ij}w_{t:t+L-1}^{il}\|Y_t^j-Y_t^l\|_2^2\\
&\geq\pi_{\min}\sum_{i=1}^N\sum_{j=1}^N\sum_{l=1}^N w_{t:t+L-1}^{ij}w_{t:t+L-1}^{il}\|Y_t^j-Y_t^l\|_2^2.
\end{aligned}$$

Let j^* and l^* be the indices such that

$$|Y_t^{j^*}-Y_t^{l^*}|=\max_{1\leq j,l\leq N}|Y_t^j-Y_t^l|.$$

From the definition of Y_t, $Y_t^j-Y_t^l=\theta_t^j-\theta_t^l$ for all $j,l\in\mathcal{V}$, which implies that

$$|Y_t^{j^*}-Y_t^{l^*}|=\max_{1\leq j,l\leq N}|Y_t^j-Y_t^l|=\max_{1\leq j,l\leq N}|\theta_t^j-\theta_t^l|=|\theta_t^{j^*}-\theta_t^{l^*}|.$$

Since $\cup_{k=t}^{t+L-1}\mathbb{G}_k$ is a strongly connected graph for all $t\geq 0$, we can find a shortest path from agent j^* to agent l^*: $(j_0,j_1),\cdots,(j_{p-1},j_p)$ with $j_0=j^*$, $j_p=l^*$, and (j_{m-1},j_m) is the edge of graph $\cup_{k=t}^{t+L-1}\mathbb{G}_k$, for $1\leq m\leq p$, which implies that

$$\begin{aligned}
&\sum_{i=1}^N\sum_{j=1}^N\sum_{l=1}^N w_{t:t+L-1}^{ij}w_{t:t+L-1}^{il}\|Y_t^j-Y_t^l\|_2^2\\
&\geq\sum_{i=1}^N\sum_{m=1}^p w_{t:t+L-1}^{ij_{m-1}}w_{t:t+L-1}^{ij_m}(Y_t^{j_{m-1}}-Y_t^{j_m})^2.
\end{aligned}\tag{34}$$

Moreover, we have

$$\sum_{i=1}^{N} w_{t:t+L-1}^{ij_{m-1}} w_{t:t+L-1}^{ij_m} \geq w_{t:t+L-1}^{j_{m-1}j_{m-1}} w_{t:t+L-1}^{j_{m-1}j_m} + w_{t:t+L-1}^{j_m j_{m-1}} w_{t:t+L-1}^{j_m j_m} \geq \beta^{2L}. \tag{35}$$

Then, from Jensen's inequality, equation 34 and equation 35, we have

$$\begin{aligned}
&\sum_{i=1}^{N} \pi_{t+L}^i \sum_{j=1}^{N}\sum_{l=1}^{N} w_{t:t+L-1}^{ij} w_{t:t+L-1}^{il} \|Y_t^j - Y_t^l\|_2^2 \\
&\geq \pi_{\min} \sum_{i=1}^{N}\sum_{m=1}^{p} w_{t:t+L-1}^{ij_{m-1}} w_{t:t+L-1}^{ij_m} (Y_t^{j_{m-1}} - Y_t^{j_m})^2 \\
&\geq \frac{\pi_{\min}\beta^{2L}}{p} (Y_t^{j^*} - Y_t^{l^*})^2 = \frac{\pi_{\min}\beta^{2L}}{\delta_t} (\theta_t^{j^*} - \theta_t^{l^*})^2.
\end{aligned} \tag{36}$$

For the case when $K > 1$, let Y_t^{ik} be the k-th entry of vector Y_t^i. Then,

$$\begin{aligned}
&\sum_{i=1}^{N} \pi_{t+L}^i \sum_{j=1}^{N}\sum_{l=1}^{N} w_{t:t+L-1}^{ij} w_{t:t+L-1}^{il} \|Y_t^j - Y_t^l\|_2^2 \\
&= \sum_{k=1}^{K}\sum_{i=1}^{N} \pi_{t+L}^i \sum_{j=1}^{N}\sum_{l=1}^{N} w_{t:t+L-1}^{ij} w_{t:t+L-1}^{il} (Y_t^{jk} - Y_t^{lk})^2.
\end{aligned}$$

For each entry k, we have

$$\sum_{i=1}^{N} \pi_{t+L}^i \sum_{j=1}^{N}\sum_{l=1}^{N} w_{t:t+L-1}^{ij} w_{t:t+L-1}^{il} (Y_t^{jk} - Y_t^{lk})^2 \geq \frac{\pi_{\min}\beta^{2L}}{\delta_{\max}} \max_{1\leq j,l\leq N} (\theta_t^{jk} - \theta_t^{lk})^2, \tag{37}$$

where θ_t^{ik} is the k-th entry of vector θ_t^i. Moreover, let $\Theta_t^{\cdot k}$ be the k-th column of matrix Θ_t.

Since $2x_1x_2 \leq x_1^2 + x_2^2$, we have for any entry $k = 1, \ldots, K$,

$$\begin{aligned}
\sum_{i=1}^{N} \pi_t^i (Y_t^{ik})^2 &= \sum_{i=1}^{N} \pi_t^i \|\theta_t^{ik} - \pi_t^\top \Theta_t^{\cdot k}\|_2^2 \\
&\leq \max_{1\leq i\leq N} \left[\theta_t^{ik} - \pi_t^\top \Theta_t^{\cdot k}\right]^2 = \max_{1\leq i\leq N} \left[\pi_t^\top (\mathbf{1}_N \theta_t^{ik} - \Theta_t^{\cdot k})\right]^2 \\
&= \max_{1\leq i\leq N} \left[\sum_{j=1}^{N} \pi_t^j (\theta_t^{ik} - \theta_t^{jk})\right]^2 = \max_{1\leq i\leq N} \sum_{j=1}^{N}\sum_{l=1}^{N} \pi_t^j \pi_t^l (\theta_t^{ik} - \theta_t^{jk})(\theta_t^{ik} - \theta_t^{lk}) \\
&\leq \max_{1\leq i\leq N} \sum_{j=1}^{N} (\pi_t^j)^2 (\theta_t^{ik} - \theta_t^{jk})^2 \leq \max_{1\leq i\leq N} \sum_{j=1}^{N} \pi_t^j (\theta_t^{ik} - \theta_t^{jk})^2 \\
&\leq \max_{1\leq i\leq N} \max_{1\leq j\leq N} (\theta_t^{ik} - \theta_t^{jk})^2.
\end{aligned}$$

Then, combining this inequality with equation 36 and equation 37, we have

$$\begin{aligned}
&\sum_{k=1}^{K}\sum_{i=1}^{N} \pi_{t+L}^i \sum_{j=1}^{N}\sum_{l=1}^{N} w_{t:t+L-1}^{ij} w_{t:t+L-1}^{il} (Y_t^{jk} - Y_t^{lk})^2 \\
&\geq \frac{\pi_{\min}\beta^{2L}}{\delta_{\max}} \sum_{k=1}^{K} \max_{1\leq j,l\leq N} (\theta_t^{jk} - \theta_t^{lk})^2 \\
&= \frac{\pi_{\min}\beta^{2L}}{\delta_{\max}} \sum_{k=1}^{K}\sum_{i=1}^{N} \pi_t^i (Y_t^{ik})^2 = \frac{\pi_{\min}\beta^{2L}}{\delta_{\max}} \sum_{i=1}^{N} \pi_t^i \|Y_t^i\|_2^2.
\end{aligned}$$

This completes the proof. ■

Lemma 5 *Suppose that Assumptions 1 and 2 hold and $\{\mathbb{G}_t\}$ is uniformly strongly connected by sub-sequences of length L. Then, when $\alpha \in (0, \zeta_1)$, we have for all $t \geq \tau(\alpha)$,*

$$\sum_{i=1}^{N} \pi_t^i \|\theta_t^i - \langle\theta\rangle_t\|_2^2 \leq \epsilon^{q_t} \sum_{i=1}^{N} \pi_{m_t}^i \|\theta_{m_t}^i - \langle\theta\rangle_{m_t}\|_2^2 + \frac{\zeta_2}{1-\epsilon},$$

where ζ_1 is defined in Appendix A, ϵ and ζ_2 are defined in equation 6 and equation 12, respectively.

Proof of Lemma 5: Let $M_t = \mathrm{diag}(\pi_t)$. Recall the update of Y_{t+L}^i,

$$Y_{t+L}^i = \left(\Pi_{k=t}^{t+L-1}(I + \alpha A(X_k))\right) \sum_{j=1}^{N} w_{t:t+L-1}^{ij} Y_t^j + \alpha \hat{b}_{t+L}^i.$$

Then, we have

$$\begin{aligned}
\|Y_{t+L}\|_{M_{t+L}}^2 &= \sum_{i=1}^{N} \pi_{t+L}^i \|Y_{t+L}^i\|_2^2 \\
&= \sum_{i=1}^{N} \pi_{t+L}^i \| \left(\Pi_{k=t}^{t+L-1}(I + \alpha A(X_k))\right) \sum_{j=1}^{N} w_{t:t+L-1}^{ij} Y_t^j \|_2^2 \qquad (38)\\
&\quad + \alpha^2 \sum_{i=1}^{N} \pi_{t+L}^i \|\hat{b}_{t+L}^i\|_2^2 \qquad (39)\\
&\quad + 2\alpha \sum_{i=1}^{N} \pi_{t+L}^i (\hat{b}_{t+L}^i)^\top \left(\Pi_{k=t}^{t+L-1}(I + \alpha A(X_k))\right) \sum_{j=1}^{N} w_{t:t+L-1}^{ij} Y_t^j. \qquad (40)
\end{aligned}$$

For equation 38, since $2(x_1)^\top x_2 = \|x_1\|_2^2 + \|x_2\|_2^2 - \|x_1 - x_2\|_2^2$ and $\pi_t^\top = \pi_{t+L}^\top W_{t:t+L-1}$, we have

$$\begin{aligned}
&\sum_{i=1}^{N} \pi_{t+L}^i \| \left(\Pi_{k=t}^{t+L-1}(I + \alpha A(X_k))\right) \sum_{j=1}^{N} w_{t:t+L-1}^{ij} Y_t^j \|_2^2 \\
&\leq (1 + \alpha A_{\max})^{2L} \sum_{i=1}^{N} \pi_{t+L}^i \| \sum_{j=1}^{N} w_{t:t+L-1}^{ij} Y_t^j \|_2^2 \\
&= (1 + \alpha A_{\max})^{2L} \sum_{i=1}^{N} \pi_{t+L}^i \sum_{j=1}^{N} \sum_{l=1}^{N} w_{t:t+L-1}^{ij} w_{t:t+L-1}^{il} \frac{1}{2} \left[\|Y_t^j\|_2^2 + \|Y_t^l\|_2^2 - \|Y_t^j - Y_t^l\|_2^2\right] \\
&= (1 + \alpha A_{\max})^{2L} \left[\sum_{i=1}^{N} \pi_t^i \|Y_t^i\|_2^2 - \frac{1}{2} \sum_{i=1}^{N} \pi_{t+L}^i \sum_{j=1}^{N} \sum_{l=1}^{N} w_{t:t+L-1}^{ij} w_{t:t+L-1}^{il} \|Y_t^j - Y_t^l\|_2^2\right].
\end{aligned}$$

From Lemma 4, we have

$$\sum_{i=1}^{N} \pi_{t+L}^i \sum_{j=1}^{N} \sum_{k=1}^{N} w_{t:t+L-1}^{ij} w_{t:t+L-1}^{ik} \|Y_t^j - Y_t^k\|_2^2 \geq \frac{\pi_{\min} \beta^{2L}}{\delta_{\max}} \sum_{i=1}^{N} \pi_t^i \|Y_t^i\|_2^2,$$

which implies that

$$\begin{aligned}
&\sum_{i=1}^{N} \pi_{t+L}^i \| \left(\Pi_{k=t}^{t+L-1}(I + \alpha A(X_k))\right) \sum_{j=1}^{N} w_{t:t+L-1}^{ij} Y_t^j \|_2^2 \\
&\leq (1 + \alpha A_{\max})^{2L} (1 - \frac{\pi_{\min} \beta^{2L}}{2\delta_{\max}}) \sum_{i=1}^{N} \pi_t^i \|Y_t^i\|_2^2. \qquad (41)
\end{aligned}$$

As for equation 39, since for any agent i we have $\|b^i(X_t) - B^\top(X_t)\pi_{t+1}\|_2 \leq 2b_{\max}$ for all i, then

$$\begin{aligned}
\|\hat{b}^i_{t+L}\|_2 &\leq \|(b^i(X_{t+L-1}) - B(X_{t+L-1})^\top \pi_{t+L})\|_2 \\
&\quad + \sum_{k=t}^{t+L-2} \|\left(\Pi_{j=k+1}^{t+L-1}(I + \alpha A(X_j))\right)\|_2 \sum_{j=1}^N w^{ij}_{k+1:t+L-1}\|(b^j(X_k) - B(X_k)^\top \pi_{k+1})\|_2 \\
&\leq 2b_{\max} \sum_{j=0}^{L-1} (1+\alpha A_{\max})^j \leq 2b_{\max}(1+\alpha A_{\max})^{L-1} \sum_{j=0}^{L-1} \frac{1}{(1+\alpha A_{\max})^j} \\
&\leq 2b_{\max} \frac{(1+\alpha A_{\max})^L - 1}{\alpha A_{\max}},
\end{aligned}$$

which implies that

$$\alpha^2 \sum_{i=1}^N \pi^i_{t+L}\|\hat{b}^i_{t+L}\|_2^2 \leq \frac{4b^2_{\max}}{A^2_{\max}} \left((1+\alpha A_{\max})^L - 1\right)^2. \tag{42}$$

In addition, since for any vector x, there holds $2\|x\|_2 \leq 1 + \|x\|_2^2$, then, for equation 40, we have

$$\begin{aligned}
&2\alpha \sum_{i=1}^N \pi^i_{t+L}(\hat{b}^i_{t+L})^\top \left(\Pi_{k=t}^{t+L-1}(I + \alpha A(X_k))\right) \sum_{j=1}^N w^{ij}_{t:t+L-1} Y^j_t \\
&\leq 2\alpha \sum_{i=1}^N \pi^i_{t+L}\|\hat{b}^i_{t+L}\|_2 \|\Pi_{k=t}^{t+L-1}(I + \alpha A(X_k))\|_2 \sum_{j=1}^N w^{ij}_{t:t+L-1}\|Y^j_t\|_2 \\
&\leq 4\alpha b_{\max} \frac{(1+\alpha A_{\max})^L - 1}{\alpha A_{\max}} (1+\alpha A_{\max})^L \sum_{i=1}^N \pi^i_t \|Y^i_t\|_2 \\
&\leq 2b_{\max} \frac{(1+\alpha A_{\max})^L - 1}{A_{\max}} (1+\alpha A_{\max})^L \left(\sum_{i=1}^N \pi^i_t \|Y^i_t\|_2^2 + 1\right).
\end{aligned} \tag{43}$$

From equation 41–equation 43, we have

$$\begin{aligned}
&\|Y_{t+L}\|^2_{M_{t+L}} \\
&\leq (1+\alpha A_{\max})^{2L}(1 - \frac{\pi_{\min}\beta^{2L}}{2\delta_{\max}}) \sum_{i=1}^N \pi^i_t \|Y^i_t\|_2^2 + \frac{4b^2_{\max}}{A^2_{\max}} \left((1+\alpha A_{\max})^L - 1\right)^2 \\
&\quad + 2b_{\max} \frac{(1+\alpha A_{\max})^L - 1}{A_{\max}} (1+\alpha A_{\max})^L \left(\sum_{i=1}^N \pi^i_t \|Y^i_t\|_2^2 + 1\right) \\
&= \left((1+\alpha A_{\max})^{2L}(1 - \frac{\pi_{\min}\beta^{2L}}{2\delta_{\max}}) + 2b_{\max} \frac{(1+\alpha A_{\max})^L - 1}{A_{\max}} (1+\alpha A_{\max})^L\right) \|Y_t\|^2_{M_t} \\
&\quad + \frac{4b^2_{\max}}{A^2_{\max}} \left((1+\alpha A_{\max})^L - 1\right)^2 + 2b_{\max} \frac{(1+\alpha A_{\max})^L - 1}{A_{\max}} (1+\alpha A_{\max})^L.
\end{aligned}$$

From Lemma 3, $0 < \epsilon < 1$ when $0 < \alpha < \zeta_1$. With the definition of ϵ and ζ_2 in equation 6 and equation 12, we have

$$\begin{aligned}
\|Y_{t+L}\|^2_{M_{t+L}} &\leq \epsilon \|Y_t\|^2_{M_t} + \zeta_2 \leq \epsilon^{q_{t+L}} \|Y_{m_t}\|^2_{M_{m_t}} + \zeta_2 \sum_{k=0}^{q_{t+L}-1} \epsilon^k \\
&\leq \epsilon^{q_{t+L}} \|Y_{m_t}\|^2_{M_{m_t}} + \frac{\zeta_2}{1-\epsilon},
\end{aligned}$$

which implies that

$$\sum_{i=1}^N \pi^i_t \|\theta^i_t - \langle\theta\rangle_t\|_2^2 \leq \epsilon^{q_t} \sum_{i=1}^N \pi^i_{m_t} \|\theta^i_{m_t} - \langle\theta\rangle_{m_t}\|_2^2 + \frac{\zeta_2}{1-\epsilon},$$

where q_t and m_t are defined in Theorem 3. This completes the proof. ■

Lemma 6 *Suppose that Assumptions 2 and 3 hold. If $\{\mathbb{G}_t\}$ is uniformly strongly connected, then when the step-size α and corresponding mixing time $\tau(\alpha)$ satisfy*

$$0 < \alpha\tau(\alpha) < \frac{\log 2}{A_{\max}},$$

we have for any $t \geq \tau(\alpha)$,

$$\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2 \leq 2\alpha A_{\max}\tau(\alpha)\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 + 2\alpha\tau(\alpha)b_{\max} \tag{44}$$

$$\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2 \leq 6\alpha\tau(\alpha)A_{\max}\|\langle\theta\rangle_t\|_2 + 5\alpha\tau(\alpha)b_{\max} \tag{45}$$

$$\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \leq 72\alpha^2\tau^2(\alpha)A_{\max}^2\|\langle\theta\rangle_t\|_2^2 + 50\alpha^2\tau^2(\alpha)b_{\max}^2 \leq 8\|\langle\theta\rangle_t\|_2^2 + \frac{6b_{\max}^2}{A_{\max}^2}. \tag{46}$$

Proof of Lemma 6: Recall the update of $\langle\theta\rangle_t$ at equation 4 with $\alpha_t = \alpha$ for all $t \geq 0$:

$$\langle\theta\rangle_{t+1} = \langle\theta\rangle_t + \alpha A(X_t)\langle\theta\rangle_t + \alpha B(X_t)^\top \pi_{t+1}.$$

Then, we have

$$\begin{aligned}
\|\langle\theta\rangle_{t+1}\|_2 &\leq \|\langle\theta\rangle_t\|_2 + \alpha A_{\max}\|\langle\theta\rangle_t\|_2 + \alpha b_{\max} \\
&\leq (1+\alpha A_{\max})\|\langle\theta\rangle_t\|_2 + \alpha b_{\max}.
\end{aligned}$$

By using $(1+x) \leq \exp(x)$, for all $u \in [t-\tau(\alpha), t]$, we have

$$\begin{aligned}
\|\langle\theta\rangle_u\|_2 &\leq (1+\alpha A_{\max})^{u-t+\tau(\alpha)}\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 + \alpha b_{\max}\sum_{l=t-\tau(\alpha)}^{u-1}(1+\alpha A_{\max})^{u-1-l} \\
&\leq (1+\alpha A_{\max})^{\tau(\alpha)}\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 + \alpha b_{\max}\sum_{l=t-\tau(\alpha)}^{u-1}(1+\alpha A_{\max})^{u-1-t+\tau(\alpha)} \\
&\leq \exp(\alpha\tau(\alpha)A_{\max})\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 + \alpha\tau(\alpha)b_{\max}\exp(\alpha\tau(\alpha)A_{\max}).
\end{aligned}$$

Since we have $\alpha\tau(\alpha)A_{\max} \leq \log 2 < \frac{1}{3}$, then $\exp(\alpha\tau(\alpha)A_{\max}) \leq 2$, which meas that

$$\|\langle\theta\rangle_u\|_2 \leq 2\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 + 2\alpha\tau(\alpha)b_{\max}.$$

Thus, we can use this to prove equation 44 for all $t \geq \tau(\alpha)$, i.e.,

$$\begin{aligned}
\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2 &\leq \sum_{u=t-\tau(\alpha)}^{t-1}\|\langle\theta\rangle_{u+1} - \langle\theta\rangle_u\|_2 \\
&\leq \alpha A_{\max}\sum_{u=t-\tau(\alpha)}^{t-1}\|\langle\theta\rangle_u\|_2 + \alpha\tau(\alpha)b_{\max} \\
&\leq \alpha A_{\max}\sum_{u=t-\tau(\alpha)}^{t-1}\left(2\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 + 2\alpha\tau(\alpha)b_{\max}\right) + \alpha\tau(\alpha)b_{\max} \\
&\leq 2\alpha\tau(\alpha)A_{\max}\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 + 2\alpha^2\tau^2(\alpha)A_{\max}b_{\max} + \alpha\tau(\alpha)b_{\max} \\
&\leq 2\alpha\tau(\alpha)A_{\max}\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 + \frac{5}{3}\alpha\tau(\alpha)b_{\max} \\
&\leq 2\alpha\tau(\alpha)A_{\max}\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 + 2\alpha\tau(\alpha)b_{\max}.
\end{aligned}$$

Moreover, we can prove equation 45 by using the equation above for all $t \geq \tau(\alpha)$ as follows:

$$\begin{aligned}
\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2 &\leq 2\alpha\tau(\alpha)A_{\max}\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 + \frac{5}{3}\alpha\tau(\alpha)b_{\max} \\
&\leq \frac{2}{3}\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2 + 2\alpha\tau(\alpha)A_{\max}\|\langle\theta\rangle_t\|_2 + \frac{5}{3}\alpha\tau(\alpha)b_{\max} \\
&\leq 6\alpha\tau(\alpha)A_{\max}\|\langle\theta\rangle_t\|_2 + 5\alpha\tau(\alpha)b_{\max}.
\end{aligned}$$

Next, using the inequality $(x+y)^2 \le 2x^2+y^2$ for all x, y, we can show equation 46 with equation 45, i.e.,

$$\begin{aligned}\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 &\le 72\alpha^2\tau^2(\alpha)A_{\max}^2\|\langle\theta\rangle_t\|_2^2 + 50\alpha^2\tau^2(\alpha)b_{\max}^2 \\ &\le 8\|\langle\theta\rangle_t\|_2^2 + \frac{6b_{\max}^2}{A_{\max}^2},\end{aligned}$$

where we use $\alpha\tau(\alpha)A_{\max} < \frac{1}{3}$ in the last inequality. ■

Lemma 7 *Let $\mathcal{F}_t = \sigma(X_k,\ k \le t)$ be a σ-algebra on $\{X_t\}$. Suppose that Assumptions 2–4 and 6 hold. If $\{\mathbb{G}_t\}$ is uniformly strongly connected, then when*

$$0 < \alpha < \frac{\log 2}{A_{\max}\tau(\alpha)},$$

we have for any $t \ge \tau(\alpha)$,

$$\begin{aligned}&\left|\mathbf{E}\left[(\langle\theta\rangle_t - \theta^*)^\top(P+P^\top)\left(A(X_t)\langle\theta\rangle_t + B(X_t)^\top\pi_{t+1} - A\langle\theta\rangle_t - b\right) \mid \mathcal{F}_{t-\tau(\alpha)}\right]\right| \\ &\le \alpha\gamma_{\max}\left(72 + 456\tau(\alpha)A_{\max}^2 + 84\tau(\alpha)A_{\max}b_{\max}\right)\mathbf{E}\left[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}\right] \\ &\quad + \alpha\gamma_{\max}\left[2 + 4\|\theta^*\|_2^2 + \frac{48b_{\max}^2}{A_{\max}^2} + \tau(\alpha)A_{\max}^2\left(152\left(\frac{b_{\max}}{A_{\max}} + \|\theta^*\|_2\right)^2\right.\right. \\ &\quad \left.\left. + \frac{48b_{\max}}{A_{\max}}\left(\frac{b_{\max}}{A_{\max}} + 1\right)^2 + \frac{87b_{\max}^2}{A_{\max}^2} + \frac{12b_{\max}}{A_{\max}}\right)\right] \\ &\quad + 2\gamma_{\max}\eta_{t+1}\sqrt{N}b_{\max}\left(1 + 9\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + \frac{6b_{\max}^2}{A_{\max}^2} + \|\theta^*\|_2^2\right).\end{aligned}$$

Proof of Lemma 7: Note that for all $t \ge \tau(\alpha)$, we have

$$\begin{aligned}&|\mathbf{E}[(\langle\theta\rangle_t - \theta^*)^\top(P+P^\top)(A(X_t)\langle\theta\rangle_t + B(X_t)^\top\pi_{t+1} - A\langle\theta\rangle_t - b) \mid \mathcal{F}_{t-\tau(\alpha)}]| \\ &\le |\mathbf{E}[(\langle\theta\rangle_t - \theta^*)^\top(P+P^\top)(A(X_t) - A)\langle\theta\rangle_t \mid \mathcal{F}_{t-\tau(\alpha)}]| \\ &\quad + |\mathbf{E}[(\langle\theta\rangle_t - \theta^*)^\top(P+P^\top)(B(X_t)^\top\pi_{t+1} - b) \mid \mathcal{F}_{t-\tau(\alpha)}]| \\ &\le |\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*)^\top(P+P^\top)(A(X_t) - A)\langle\theta\rangle_{t-\tau(\alpha)} \mid \mathcal{F}_{t-\tau(\alpha)}]| \qquad (47) \\ &\quad + |\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*)^\top(P+P^\top)(A(X_t) - A)(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}) \mid \mathcal{F}_{t-\tau(\alpha)}]| \qquad (48) \\ &\quad + |\mathbf{E}[(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)})^\top(P+P^\top)(A(X_t) - A)\langle\theta\rangle_{t-\tau(\alpha)} \mid \mathcal{F}_{t-\tau(\alpha)}]| \qquad (49) \\ &\quad + |\mathbf{E}[(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)})^\top(P+P^\top)(A(X_t) - A)(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}) \mid \mathcal{F}_{t-\tau(\alpha)}]| \qquad (50) \\ &\quad + |\mathbf{E}[(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)})^\top(P+P^\top)(B(X_t)^\top\pi_{t+1} - b) \mid \mathcal{F}_{t-\tau(\alpha)}]| \qquad (51) \\ &\quad + |\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*)^\top(P+P^\top)(B(X_t)^\top\pi_{t+1} - b) \mid \mathcal{F}_{t-\tau(\alpha)}]|. \qquad (52)\end{aligned}$$

First, by using the mixing time in Assumption 3, we can get the bound for equation 47 and equation 52 for all $t \ge \tau(\alpha)$ as follows:

$$\begin{aligned}&|\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*)^\top(P+P^\top)(A(X_t) - A)\langle\theta\rangle_{t-\tau(\alpha)} \mid \mathcal{F}_{t-\tau(\alpha)}]| \\ &\le |(\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*)^\top(P+P^\top)\mathbf{E}[A(X_t) - A \mid \mathcal{F}_{t-\tau(\alpha)}]\langle\theta\rangle_{t-\tau(\alpha)}| \\ &\le 2\alpha\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*\|_2\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\ &\le \alpha\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*\|_2^2 + \|\langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\ &\le \alpha\gamma_{\max}\mathbf{E}[2\|\theta^*\|_2^2 + 3\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\ &\le 6\alpha\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 6\alpha\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 2\alpha\gamma_{\max}\|\theta^*\|_2^2 \\ &\le 54\alpha\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 36\alpha\gamma_{\max}(\frac{b_{\max}}{A_{\max}})^2 + 2\alpha\gamma_{\max}\|\theta^*\|_2^2, \qquad (53)\end{aligned}$$

where in the last inequality, we use equation 44 from Lemma 6. Then, from the definition of π_∞ in Assumption 6,

$$
\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*)^\top (P + P^\top)(B(X_t)^\top \pi_{t+1} - b) \mid \mathcal{F}_{t-\tau(\alpha)}]| \\
&\leq |\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*)^\top (P + P^\top)(\sum_{i=1}^N \pi_{t+1}^i (b^i(X_t) - b^i) + \sum_{i=1}^N (\pi_{t+1}^i - \pi_\infty^i) b^i) \mid \mathcal{F}_{t-\tau(\alpha)}]| \\
&\leq |(\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*)^\top (P + P^\top)(\sum_{i=1}^N \pi_{t+1}^i \mathbf{E}[b^i(X_t) - b^i \mid \mathcal{F}_{t-\tau(\alpha)}] + \sum_{i=1}^N (\pi_{t+1}^i - \pi_\infty^i) b^i)| \\
&\leq 2\gamma_{\max}(\alpha + \eta_{t+1}\sqrt{N} b_{\max}) \mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*\|_2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\leq 2\gamma_{\max}(\alpha + \eta_{t+1}\sqrt{N} b_{\max}) \left(\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha)}] + \|\theta^*\|_2\right) \\
&\leq 2\gamma_{\max}(\alpha + \eta_{t+1}\sqrt{N} b_{\max}) \left(1 + \frac{1}{2}\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + \frac{1}{2}\|\theta^*\|_2^2\right) \\
&\leq 2\gamma_{\max}(\alpha + \eta_{t+1}\sqrt{N} b_{\max}) \left(1 + \mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 + \|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + \|\theta^*\|_2^2\right) \\
&\leq 2\gamma_{\max}(\alpha + \eta_{t+1}\sqrt{N} b_{\max}) \left(1 + 9\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 6(\frac{b_{\max}}{A_{\max}})^2 + \|\theta^*\|_2^2\right), \qquad (54)
\end{aligned}
$$

where we also use equation 44 from Lemma 6 in the last inequality.

Next, by using Assumption 2, equation 44 and equation 46, we have

$$
\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*)^\top (P + P^\top)(A(X_t) - A)(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}) \mid \mathcal{F}_{t-\tau(\alpha)}]| \\
&\leq 4\gamma_{\max} A_{\max} \mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)} - \theta^*\|_2 \|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\leq 4\gamma_{\max} A_{\max} \mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 \|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\quad + 4\gamma_{\max} A_{\max} \|\theta^*\|_2 \mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\leq 8\alpha\tau(\alpha)\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 8\alpha\tau(\alpha)\gamma_{\max} A_{\max} b_{\max} \|\theta^*\|_2 \\
&\quad + 8\alpha\tau(\alpha)\gamma_{\max} A_{\max}^2 \left(\frac{b_{\max}}{A_{\max}} + \|\theta^*\|_2\right) \mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\leq 8\alpha\tau(\alpha)\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 8\alpha\tau(\alpha)\gamma_{\max} A_{\max} b_{\max} \|\theta^*\|_2 \\
&\quad + 4\alpha\tau(\alpha)\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 4\alpha\tau(\alpha)\gamma_{\max} A_{\max}^2 \left(\frac{b_{\max}}{A_{\max}} + \|\theta^*\|_2\right)^2 \\
&\leq 12\alpha\tau(\alpha)\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 8\alpha\tau(\alpha)\gamma_{\max} \left(b_{\max} + A_{\max}\|\theta^*\|_2\right)^2 \\
&\leq 24\alpha\tau(\alpha)\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 8\alpha\tau(\alpha)\gamma_{\max} \left(b_{\max} + A_{\max}\|\theta^*\|_2\right)^2 \\
&\quad + 24\alpha\tau(\alpha)\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\leq 216\alpha\tau(\alpha)\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 144\alpha\tau(\alpha)\gamma_{\max} b_{\max}^2 \\
&\quad + 8\alpha\tau(\alpha)\gamma_{\max} \left(b_{\max} + A_{\max}\|\theta^*\|_2\right)^2 \\
&\leq 216\alpha\tau(\alpha)\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 152\alpha\tau(\alpha)\gamma_{\max} \left(b_{\max} + A_{\max}\|\theta^*\|_2\right)^2. \qquad (55)
\end{aligned}
$$

In additional, by using equation 44 and equation 46, we have

$$
\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)})^\top (P + P^\top)(A(X_t) - A)\langle\theta\rangle_{t-\tau(\alpha)} \mid \mathcal{F}_{t-\tau(\alpha)}]| \\
&\leq 4\gamma_{\max} A_{\max} \mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2 \|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\leq 8\alpha\tau(\alpha)\gamma_{\max} A_{\max} \mathbf{E}[A_{\max}\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 + b_{\max}\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\leq 4\alpha\tau(\alpha)\gamma_{\max} A_{\max}(2A_{\max} + b_{\max}) \mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 4\alpha\tau(\alpha)\gamma_{\max} A_{\max} b_{\max} \\
&\leq 8\alpha\tau(\alpha)\gamma_{\max} A_{\max}(2A_{\max} + b_{\max}) \mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\quad + 8\alpha\tau(\alpha)\gamma_{\max} A_{\max}(2A_{\max} + b_{\max}) \mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 4\alpha\tau(\alpha)\gamma_{\max} A_{\max} b_{\max} \\
&\leq 72\alpha\tau(\alpha)\gamma_{\max} A_{\max}(2A_{\max} + b_{\max}) \mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\quad + 48\alpha\tau(\alpha)\gamma_{\max} A_{\max} b_{\max} (\frac{b_{\max}}{A_{\max}} + 1)^2. \qquad (56)
\end{aligned}
$$

Moreover, we can get the bound for equation 50 by using equation 46 as follows:

$$
\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)})^\top (P+P^\top)(A(X_t)-A)(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}) \mid \mathcal{F}_{t-\tau(\alpha)}]| \\
&\leq 4\gamma_{\max} A_{\max} \mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}]| \\
&\leq 4\gamma_{\max} A_{\max} \mathbf{E}[72\alpha^2\tau^2(\alpha)A_{\max}^2 \|\langle\theta\rangle_t\|_2^2 + 50\alpha^2\tau^2(\alpha)b_{\max}^2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\leq 96\alpha\tau(\alpha)A_{\max}^2\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 67\alpha\tau(\alpha)b_{\max}^2\gamma_{\max}.
\end{aligned} \tag{57}
$$

Finally, using equation 45 we can get the bound for equation 51:

$$
\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)})^\top (P+P^\top)(B(X_t)^\top\pi_{t+1} - b) \mid \mathcal{F}_{t-\tau(\alpha)}]| \\
&\leq 4\gamma_{\max} b_{\max} \mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\leq 4\gamma_{\max} b_{\max} \mathbf{E}[6\alpha\tau(\alpha)A_{\max}\|\langle\theta\rangle_t\|_2 + 5\alpha\tau(\alpha)b_{\max} \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\leq 12\alpha\tau(\alpha)\gamma_{\max}A_{\max}b_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 12\alpha\tau(\alpha)\gamma_{\max}A_{\max}b_{\max} + 20\alpha\tau(\alpha)b_{\max}^2\gamma_{\max}.
\end{aligned} \tag{58}
$$

Then, by using equation 53–equation 58, we have

$$
\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_t^\top - \theta^*)^\top (P+P^\top)(A(X_t)\langle\theta\rangle_t + B(X_t)^\top\pi_{t+1} - A\langle\theta\rangle_t - b) \mid \mathcal{F}_{t-\tau(\alpha)}]| \\
&\leq 54\alpha\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 36\alpha\gamma_{\max}(\frac{b_{\max}}{A_{\max}})^2 + 2\alpha\gamma_{\max}\|\theta^*\|_2^2 \\
&\quad + 216\alpha\tau(\alpha)\gamma_{\max}A_{\max}^2\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 152\alpha\tau(\alpha)\gamma_{\max}\left(b_{\max} + A_{\max}\|\theta^*\|_2\right)^2 \\
&\quad + 72\alpha\tau(\alpha)\gamma_{\max}A_{\max}(2A_{\max} + b_{\max})\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\quad + 48\alpha\tau(\alpha)\gamma_{\max}A_{\max}b_{\max}(\frac{b_{\max}}{A_{\max}} + 1)^2 + 20\alpha\tau(\alpha)b_{\max}^2\gamma_{\max} \\
&\quad + 96\alpha\tau(\alpha)A_{\max}^2\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 67\alpha\tau(\alpha)b_{\max}^2\gamma_{\max} \\
&\quad + 12\alpha\tau(\alpha)\gamma_{\max}A_{\max}b_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 12\alpha\tau(\alpha)\gamma_{\max}A_{\max}b_{\max} \\
&\quad + 2\gamma_{\max}(\alpha + \eta_{t+1}\sqrt{N}b_{\max})\left(1 + 9\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 6(\frac{b_{\max}}{A_{\max}})^2 + \|\theta^*\|_2^2\right) \\
&\leq \alpha\gamma_{\max}\left(72 + 456\tau(\alpha)A_{\max}^2 + 84\tau(\alpha)A_{\max}b_{\max}\right)\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] \\
&\quad + \alpha\gamma_{\max}\Big[2 + 4\|\theta^*\|_2^2 + 48(\frac{b_{\max}}{A_{\max}})^2 + \tau(\alpha)A_{\max}^2\Big(152\left(\frac{b_{\max}}{A_{\max}} + \|\theta^*\|_2\right)^2 \\
&\quad + 48\frac{b_{\max}}{A_{\max}}(\frac{b_{\max}}{A_{\max}} + 1)^2 + 87(\frac{b_{\max}}{A_{\max}})^2 + 12\frac{b_{\max}}{A_{\max}}\Big)\Big] \\
&\quad + 2\gamma_{\max}\eta_{t+1}\sqrt{N}b_{\max}\left(1 + 9\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}] + 6(\frac{b_{\max}}{A_{\max}})^2 + \|\theta^*\|_2^2\right).
\end{aligned} \tag{59}
$$

This completes the proof. ■

Lemma 8 *Suppose that Assumptions 2–4 and 6 hold. Then, when*

$$
0 < \alpha < \min\left\{\frac{\log 2}{A_{\max}\tau(\alpha)},\ \frac{0.1}{K_2\gamma_{\max}}\right\},
$$

we have for any $t \geq T_1$,

$$
\begin{aligned}
\mathbf{E}[\|\langle\theta\rangle_{t+1} - \theta^*\|_2^2] \leq & \left(1 - \frac{0.9\alpha}{\gamma_{\max}}\right)^{t-T_1} \frac{\gamma_{\max}}{\gamma_{\min}} \mathbf{E}\left[\|\langle\theta\rangle_{T_1} - \theta^*\|_2^2\right] + \frac{\alpha\zeta_3\gamma_{\max}^2}{0.9\gamma_{\min}} \\
& + \frac{\gamma_{\max}}{\gamma_{\min}} \alpha\zeta_4 \sum_{k=0}^{t-T_1} \eta_{t+1-k} (1 - \frac{0.9\alpha}{\gamma_{\max}})^k \\
\leq & \left(1 - \frac{0.9\alpha}{\gamma_{\max}}\right)^{t+1-T_1} \frac{\gamma_{\max}}{\gamma_{\min}} (4\exp\{2\alpha A_{\max} T_1\} + 2) \mathbf{E}[\|\langle\theta\rangle_0 - \theta^*\|_2^2] \\
& + 4\left(1 - \frac{0.9\alpha}{\gamma_{\max}}\right)^{t+1-T_1} \frac{\gamma_{\max}}{\gamma_{\min}} \exp\{2\alpha A_{\max} T_1\} (\|\theta^*\|_2 + \frac{b_{\max}}{A_{\max}})^2 \\
& + \frac{\alpha\zeta_3\gamma_{\max}^2}{0.9\gamma_{\min}} + \frac{\gamma_{\max}}{\gamma_{\min}} \alpha\zeta_4 \sum_{k=0}^{t-T_1} \eta_{t+1-k} (1 - \frac{0.9\alpha}{\gamma_{\max}})^k.
\end{aligned}
$$

where ζ_3, ζ_4 and K_2 are defined inequation 13, equation 14 and equation 20, respectively.

Proof of Lemma 8: Let $H(\langle\theta\rangle_t) = (\langle\theta\rangle_t - \theta^*)^\top P(\langle\theta\rangle_t - \theta^*)$. From Assumption 4, we know that

$$
\gamma_{\min}\|\langle\theta\rangle_t - \theta^*\|_2^2 \leq H(\langle\theta\rangle_t) \leq \gamma_{\max}\|\langle\theta\rangle_t - \theta^*\|_2^2.
$$

Moreover, from Assumption 2, for all $t \geq 0$ we have

$$
\begin{aligned}
& H(\langle\theta\rangle_{t+1}) \\
&= (\langle\theta\rangle_{t+1} - \theta^*)^\top P(\langle\theta\rangle_{t+1} - \theta^*) \\
&= (\langle\theta\rangle_t + \alpha A(X_t)\langle\theta\rangle_t + \alpha B(X_t)^\top \pi_{t+1} - \theta^*)^\top P(\langle\theta\rangle_t + \alpha A(X_t)\langle\theta\rangle_t + \alpha B(X_t)^\top \pi_{t+1} - \theta^*) \\
&= (\langle\theta\rangle_t - \theta^*)^\top P(\langle\theta\rangle_t - \theta^*) + \alpha^2 (A(X_t)\langle\theta\rangle_t)^\top P(A(X_t)\langle\theta\rangle_t) \\
&\quad + \alpha^2 (B(X_t)^\top \pi_{t+1})^\top P(B(X_t)^\top \pi_{t+1}) + \alpha^2 (A(X_t)\langle\theta\rangle_t)^\top (P + P^\top)(B(X_t)^\top \pi_{t+1}) \\
&\quad + \alpha(\langle\theta\rangle_t - \theta^*)^\top (P + P^\top)(A(X_t)\langle\theta\rangle_t + B(X_t)^\top \pi_{t+1} - A\langle\theta\rangle_t - b) \\
&\quad + \alpha(\langle\theta\rangle_t - \theta^*)^\top P(A\langle\theta\rangle_t + b) + \alpha(A\langle\theta\rangle_t + b)^\top P(\langle\theta\rangle_t - \theta^*) \\
&= H(\langle\theta\rangle_t) + \alpha^2 (A(X_t)\langle\theta\rangle_t)^\top P(A(X_t)\langle\theta\rangle_t) \\
&\quad + \alpha^2 (B(X_t)^\top \pi_{t+1})^\top P(B(X_t)^\top \pi_{t+1}) + \alpha^2 (A(X_t)\langle\theta\rangle_t)^\top (P + P^\top)(B(X_t)^\top \pi_{t+1}) \\
&\quad + \alpha(\langle\theta\rangle_t - \theta^*)^\top (P + P^\top)(A(X_t)\langle\theta\rangle_t + B(X_t)^\top \pi_{t+1} - A\langle\theta\rangle_t - b) \\
&\quad + \alpha(\langle\theta\rangle_t - \theta^*)^\top (PA + A^\top P)(\langle\theta\rangle_t - \theta^*),
\end{aligned} \tag{60}
$$

where we use the fact that $A\theta^* + b = 0$ on the last equality.

Next, we can take expectation on both sides of equation 60. From Assumption 4 and Lemma 7, for $t \geq T_1$ we have

$$
\begin{aligned}
&\mathbf{E}[H(\langle\theta\rangle_{t+1})] \\
&= \mathbf{E}[H(\langle\theta\rangle_t)] + \alpha^2\mathbf{E}[(A(X_t)\langle\theta\rangle_t)^\top P(A(X_t)\langle\theta\rangle_t)] - \alpha\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] \\
&\quad + \alpha^2\mathbf{E}[(B(X_t)^\top\pi_{t+1})^\top P(B(X_t)^\top\pi_{t+1})] + \alpha^2\mathbf{E}[(A(X_t)\langle\theta\rangle_t)^\top(P+P^\top)(B(X_t)^\top\pi_{t+1})] \\
&\quad + \alpha\mathbf{E}[(\langle\theta\rangle_t - \theta^*)^\top(P+P^\top)(A(X_t)\langle\theta\rangle_t + B(X_t)^\top\pi_{t+1} - A\langle\theta\rangle_t - b)] \\
&\leq \mathbf{E}[H(\langle\theta\rangle_t)] - \alpha\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] + \alpha^2A_{\max}^2\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2^2] + 2\alpha^2A_{\max}b_{\max}\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2] \\
&\quad + \alpha^2b_{\max}^2\gamma_{\max} + \alpha^2\gamma_{\max}\left(72 + 456\tau(\alpha)A_{\max}^2 + 84\tau(\alpha)A_{\max}b_{\max}\right)\mathbf{E}[\|\langle\theta\rangle_t\|_2^2] \\
&\quad + \alpha^2\gamma_{\max}\Big[2 + 4\|\theta^*\|_2^2 + 48(\frac{b_{\max}}{A_{\max}})^2 + \tau(\alpha)A_{\max}^2\Big(152\left(\frac{b_{\max}}{A_{\max}} + \|\theta^*\|_2\right)^2 \\
&\quad + 48\frac{b_{\max}}{A_{\max}}(\frac{b_{\max}}{A_{\max}}+1)^2 + 87(\frac{b_{\max}}{A_{\max}})^2 + 12\frac{b_{\max}}{A_{\max}}\Big)\Big] \\
&\quad + 2\alpha\gamma_{\max}\eta_{t+1}\sqrt{N}b_{\max}\left(1 + 9\mathbf{E}[\|\langle\theta\rangle_t\|_2^2] + 6(\frac{b_{\max}}{A_{\max}})^2 + \|\theta^*\|_2^2\right) \\
&\leq \mathbf{E}[H(\langle\theta\rangle_t)] - \alpha\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] + 2\alpha^2A_{\max}^2\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2^2] + 2\alpha^2b_{\max}^2\gamma_{\max} \\
&\quad + \alpha^2\gamma_{\max}\left(72 + 456\tau(\alpha)A_{\max}^2 + 84\tau(\alpha)A_{\max}b_{\max}\right)\mathbf{E}[\|\langle\theta\rangle_t\|_2^2] \\
&\quad + \alpha^2\gamma_{\max}\Big[2 + 4\|\theta^*\|_2^2 + 48(\frac{b_{\max}}{A_{\max}})^2 + \tau(\alpha)A_{\max}^2\Big(152\left(\frac{b_{\max}}{A_{\max}} + \|\theta^*\|_2\right)^2 \\
&\quad + 48\frac{b_{\max}}{A_{\max}}(\frac{b_{\max}}{A_{\max}}+1)^2 + 87(\frac{b_{\max}}{A_{\max}})^2 + 12\frac{b_{\max}}{A_{\max}}\Big)\Big] \\
&\quad + 2\alpha\gamma_{\max}\eta_{t+1}\sqrt{N}b_{\max}\left(1 + 9\mathbf{E}[\|\langle\theta\rangle_t\|_2] + 6(\frac{b_{\max}}{A_{\max}})^2 + \|\theta^*\|_2^2\right).
\end{aligned}
$$

Since $\mathbf{E}[\|\langle\theta\rangle_t\|_2^2] \leq 2\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] + 2\|\theta^*\|_2^2$, we have

$$
\begin{aligned}
&\mathbf{E}[H(\langle\theta\rangle_{t+1})] \\
&\leq \mathbf{E}[H(\langle\theta\rangle_t)] - \alpha\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] + 2\alpha^2b_{\max}^2\gamma_{\max} \\
&\quad + \alpha^2\gamma_{\max}\left(72 + 2A_{\max}^2 + 456\tau(\alpha)A_{\max}^2 + 84\tau(\alpha)A_{\max}b_{\max}\right)(2\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] + 2\|\theta^*\|_2^2) \\
&\quad + \alpha^2\gamma_{\max}\Big[2 + 4\|\theta^*\|_2^2 + 48(\frac{b_{\max}}{A_{\max}})^2 + \tau(\alpha)A_{\max}^2\Big(152\left(\frac{b_{\max}}{A_{\max}} + \|\theta^*\|_2\right)^2 \\
&\quad + 48\frac{b_{\max}}{A_{\max}}(\frac{b_{\max}}{A_{\max}}+1)^2 + 87(\frac{b_{\max}}{A_{\max}})^2 + 12\frac{b_{\max}}{A_{\max}}\Big)\Big] \\
&\quad + 2\alpha\gamma_{\max}\eta_{t+1}\sqrt{N}b_{\max}\left(1 + 18\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] + 6(\frac{b_{\max}}{A_{\max}})^2 + 19\|\theta^*\|_2^2\right) \\
&\leq \mathbf{E}[H(\langle\theta\rangle_t)] \\
&\quad + \left(-\alpha + 2\alpha^2\gamma_{\max}\left(72 + 2A_{\max}^2 + 456\tau(\alpha)A_{\max}^2 + 84\tau(\alpha)A_{\max}b_{\max}\right)\right)\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] \\
&\quad + 2\alpha^2\gamma_{\max}\left(72 + 2A_{\max}^2 + 456\tau(\alpha)A_{\max}^2 + 84\tau(\alpha)A_{\max}b_{\max}\right)\|\theta^*\|_2^2 \\
&\quad + \alpha^2\gamma_{\max}\Big[2 + 2b_{\max}^2 + 4\|\theta^*\|_2^2 + 48(\frac{b_{\max}}{A_{\max}})^2 + \tau(\alpha)A_{\max}^2\Big(152\left(\frac{b_{\max}}{A_{\max}} + \|\theta^*\|_2\right)^2 \\
&\quad + 48\frac{b_{\max}}{A_{\max}}(\frac{b_{\max}}{A_{\max}}+1)^2 + 87(\frac{b_{\max}}{A_{\max}})^2 + 12\frac{b_{\max}}{A_{\max}}\Big)\Big] \\
&\quad + 2\alpha\gamma_{\max}\eta_{t+1}\sqrt{N}b_{\max}\left(1 + 18\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] + 6(\frac{b_{\max}}{A_{\max}})^2 + 19\|\theta^*\|_2^2\right) \\
&\leq \mathbf{E}[H(\langle\theta\rangle_t)] + \left(-\alpha + \alpha^2\gamma_{\max}K_2 + 36\alpha\eta_{t+1}\sqrt{N}b_{\max}\gamma_{\max}\right)\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] \\
&\quad + \alpha^2\zeta_3\gamma_{\max} + \alpha\gamma_{\max}\eta_{t+1}\zeta_4.
\end{aligned}
$$

From Lemma 3, $1 - \frac{0.9\alpha}{\gamma_{\max}} \in (0,1)$. In addition, from the definition of T_1 and $\alpha < \frac{0.1}{K_2\gamma_{\max}}$, we have

$$
\begin{aligned}
&\mathbf{E}[H(\langle\theta\rangle_{t+1})] \\
&\leq \mathbf{E}[H(\langle\theta\rangle_t)] - 0.9\alpha\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] + \alpha^2\zeta_3\gamma_{\max} + \alpha\gamma_{\max}\eta_{t+1}\zeta_4 \\
&\leq \left(1 - \frac{0.9\alpha}{\gamma_{\max}}\right)\mathbf{E}[H(\langle\theta\rangle_t)] + \alpha^2\zeta_3\gamma_{\max} + \alpha\gamma_{\max}\eta_{t+1}\zeta_4 \\
&\leq \left(1 - \frac{0.9\alpha}{\gamma_{\max}}\right)^{t+1-T_1}\mathbf{E}[H(\langle\theta\rangle_{T_1})] + \alpha^2\zeta_3\gamma_{\max}\sum_{k=T_1}^{t}(1 - \frac{0.9\alpha}{\gamma_{\max}})^{t-k} \\
&\quad + \alpha\gamma_{\max}\zeta_4\sum_{k=0}^{t-T_1}\eta_{t+1-k}(1 - \frac{0.9\alpha}{\gamma_{\max}})^k \\
&\leq \left(1 - \frac{0.9\alpha}{\gamma_{\max}}\right)^{t+1-T_1}\mathbf{E}[H(\langle\theta\rangle_{T_1})] + \frac{\alpha\zeta_3\gamma_{\max}^2}{0.9} + \alpha\gamma_{\max}\zeta_4\sum_{k=0}^{t-T_1}\eta_{t+1-k}(1 - \frac{0.9\alpha}{\gamma_{\max}})^k, \quad (61)
\end{aligned}
$$

which implies that

$$
\begin{aligned}
&\mathbf{E}[\|\langle\theta\rangle_{t+1} - \theta^*\|_2^2] \\
&\leq \frac{1}{\gamma_{\min}}\mathbf{E}[H(\langle\theta\rangle_{t+1})] \\
&\leq \left(1 - \frac{0.9\alpha}{\gamma_{\max}}\right)^{t+1-T_1}\frac{\gamma_{\max}}{\gamma_{\min}}\mathbf{E}[\|\langle\theta\rangle_{T_1} - \theta^*\|_2^2] + \frac{\alpha\zeta_3\gamma_{\max}^2}{0.9\gamma_{\min}} \\
&\quad + \frac{\gamma_{\max}}{\gamma_{\min}}\alpha\zeta_4\sum_{k=0}^{t-T_1}\eta_{t+1-k}(1 - \frac{0.9\alpha}{\gamma_{\max}})^k. \quad (62)
\end{aligned}
$$

Next, we consider the bound for $\mathbf{E}[\|\langle\theta\rangle_{T_1} - \theta^*\|_2^2]$. Since $1 + x \leq \exp\{x\}$ for any x, we have for any t,

$$
\begin{aligned}
&\|\langle\theta\rangle_{t+1} - \langle\theta\rangle_0\|_2 \\
&= \|\langle\theta\rangle_t - \langle\theta\rangle_0 + \alpha A(X_t)(\langle\theta\rangle_t - \langle\theta\rangle_0) + \alpha B(X_t)^\top\pi_{t+1} + \alpha A(X_t)\langle\theta\rangle_0\|_2 \\
&\leq (1 + \alpha A_{\max})\|\langle\theta\rangle_t - \langle\theta\rangle_0\|_2 + \alpha\left(A_{\max}\|\langle\theta\rangle_0\|_2 + b_{\max}\right) \\
&\leq \alpha\left(A_{\max}\|\langle\theta\rangle_0\|_2 + b_{\max}\right)\sum_{l=0}^{t}(1 + \alpha A_{\max})^l \\
&\leq \left(A_{\max}\|\langle\theta\rangle_0\|_2 + b_{\max}\right)\frac{(1 + \alpha A_{\max})^{t+1}}{A_{\max}} \\
&\leq \left(\|\langle\theta\rangle_0 - \theta^*\|_2 + \|\theta^*\|_2 + \frac{b_{\max}}{A_{\max}}\right)\exp\left\{\alpha A_{\max}(t+1)\right\},
\end{aligned}
$$

which implies that

$$
\|\langle\theta\rangle_{T_1} - \langle\theta\rangle_0\|_2 \leq \left(\|\langle\theta\rangle_0 - \theta^*\|_2 + \|\theta^*\|_2 + \frac{b_{\max}}{A_{\max}}\right)\exp\left\{\alpha A_{\max}T_1\right\}.
$$

Then, we have

$$
\begin{aligned}
\mathbf{E}[\|\langle\theta\rangle_{T_1} - \theta^*\|_2^2] &\leq 2\|\langle\theta\rangle_{T_1} - \langle\theta\rangle_0\|_2^2 + 2\|\langle\theta\rangle_0 - \theta^*\|_2^2 \\
&\leq (4\exp\left\{2\alpha A_{\max}T_1\right\} + 2)\mathbf{E}[\|\langle\theta\rangle_0 - \theta^*\|_2^2] \\
&\quad + 4\exp\left\{2\alpha A_{\max}T_1\right\}(\|\theta^*\|_2 + \frac{b_{\max}}{A_{\max}})^2. \quad (63)
\end{aligned}
$$

From equation 62 and equation 63, we have

$$
\begin{aligned}
\mathbf{E}[\|\langle\theta\rangle_{t+1}-\theta^*\|_2^2] \le& \left(1-\frac{0.9\alpha}{\gamma_{\max}}\right)^{t+1-T_1}\frac{\gamma_{\max}}{\gamma_{\min}}(4\exp\{2\alpha A_{\max}T_1\}+2)\mathbf{E}[\|\langle\theta\rangle_0-\theta^*\|_2^2] \\
&+4\left(1-\frac{0.9\alpha}{\gamma_{\max}}\right)^{t+1-T_1}\frac{\gamma_{\max}}{\gamma_{\min}}\exp\{2\alpha A_{\max}T_1\}(\|\theta^*\|_2+\frac{b_{\max}}{A_{\max}})^2 \\
&+\frac{\alpha\zeta_3\gamma_{\max}^2}{0.9\gamma_{\min}}+\frac{\gamma_{\max}}{\gamma_{\min}}\alpha\zeta_4\sum_{k=0}^{t-T_1}\eta_{t+1-k}(1-\frac{0.9\alpha}{\gamma_{\max}})^k.
\end{aligned}
$$

This completes the proof. ■

We are now in a position to prove the fixed step-size case in Theorem 3.

Proof of Case 1) in Theorem 3: From Lemmas 5 and 8, for any $t \ge T_1$, we have

$$
\begin{aligned}
\sum_{i=1}^N \pi_t^i\mathbf{E}[\|\theta_t^i-\theta^*\|_2^2] \le& 2\sum_{i=1}^N \pi_t^i\mathbf{E}[\|\theta_t^i-\langle\theta\rangle_t\|_2^2]+2\mathbf{E}[\|\langle\theta\rangle_t-\theta^*\|_2^2] \\
\le& 2\epsilon^{q_t}\sum_{i=1}^N \pi_{m_t}^i\mathbf{E}[\|\theta_{m_t}^i-\langle\theta\rangle_{m_t}\|_2^2]+\frac{2\zeta_2}{1-\epsilon}+\frac{2\alpha\zeta_3\gamma_{\max}^2}{0.9\gamma_{\min}} \\
&+\frac{\gamma_{\max}}{\gamma_{\min}}2\alpha\zeta_4\sum_{k=0}^{t-T_1}\eta_{t+1-k}(1-\frac{0.9\alpha}{\gamma_{\max}})^k \\
&+\left(1-\frac{0.9\alpha}{\gamma_{\max}}\right)^{t-T_1}\frac{\gamma_{\max}}{\gamma_{\min}}(8\exp\{2\alpha A_{\max}T_1\}+4)\mathbf{E}[\|\langle\theta\rangle_0-\theta^*\|_2^2] \\
&+8\left(1-\frac{0.9\alpha}{\gamma_{\max}}\right)^{t-T_1}\frac{\gamma_{\max}}{\gamma_{\min}}\exp\{2\alpha A_{\max}T_1\}(\|\theta^*\|_2+\frac{b_{\max}}{A_{\max}})^2 \\
\le& 2\epsilon^{q_t}\sum_{i=1}^N \pi_{m_t}^i\mathbf{E}\left[\left\|\theta_{m_t}^i-\langle\theta\rangle_{m_t}\right\|_2^2\right]+C_1\left(1-\frac{0.9\alpha}{\gamma_{\max}}\right)^{t-T_1}+C_2 \\
&+\frac{\gamma_{\max}}{\gamma_{\min}}2\alpha\zeta_4\sum_{k=0}^{t-T_1}\eta_{t+1-k}(1-\frac{0.9\alpha}{\gamma_{\max}})^k,
\end{aligned}
$$

where C_1 and C_2 are defined in Appendix A.1. This completes the proof. ■

E.2.2 TIME-VARYING STEP-SIZE

In this subsection, we consider the time-varying step-size case and begin with a property of η_t.

Lemma 9 *Suppose that Assumption 6 holds. Then,* $\lim_{t\to\infty}\eta_t = 0$ *and* $\lim_{t\to\infty}\frac{1}{t+1}\sum_{k=0}^t \eta_k = 0$.

Proof of Lemma 9: From Assumption 6, we know that π_t will converge to π_∞, and thus η_t will converge to 0. Next, we will prove that $\lim_{t\to\infty}\frac{1}{t+1}\sum_{k=0}^t\eta_k = 0$. For any positive constant $c > 0$, there exists a positive integer $T(c)$, depending on c, such that $\forall t \ge T(c)$, we have $\eta_t < c$. Thus,

$$
\frac{1}{t}\sum_{k=0}^{t-1}\eta_k = \frac{1}{t}\sum_{k=0}^{T(c)}\eta_k+\frac{1}{t}\sum_{k=T(c)+1}^{t-1}\eta_k \le \frac{1}{t}\sum_{k=0}^{T(c)}\eta_k+\frac{t-1-T(c)}{t}c.
$$

Let $t\to\infty$ on both sides of the above inequality. Then, we have

$$
\lim_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t-1}\eta_k \le \lim_{t\to\infty}\frac{1}{t}\sum_{k=0}^{T(c)}\eta_k+\lim_{t\to\infty}\frac{t-1-T(c)}{t}c = c.
$$

Since the above argument holds for arbitrary positive c, then $\lim_{t\to\infty} \frac{1}{t+1}\sum_{k=0}^{t} \eta_k = 0$. ■

Recall the updates corresponding to the time-varying step-size case given in equation 3 and equation 4,

$$\Theta_{t+1} = W_t\Theta_t + \alpha_t W_t \Theta_t A(X_t)^\top + \alpha_t B(X_t),$$
$$\langle\theta\rangle_{t+1} = \langle\theta\rangle_t + \alpha_t A(X_t)\langle\theta\rangle_t + \alpha_t B(X_t)^\top \pi_{t+1}.$$

From equation 33, we get the update for Y_t with the time-varying step-size as follows:

$$\begin{aligned} Y_{t+L} = {} & W_{t:t+L-1} Y_t (I + \alpha_t A^\top(X_t)) \cdots (I + \alpha_{t+L-1} A^\top (X_{t+L-1})) \\ & + \alpha_{t+L-1}(I - \mathbf{1}_N \pi_{t+L}^\top) B(X_{t+L-1}) \\ & + \sum_{k=t}^{t+L-2} \alpha_k W_{k+1:t+L-1}(I - \mathbf{1}_N \pi_{k+1}^\top) B(X_k) \left(\Pi_{j=k+1}^{t+L-1} (I + \alpha_j A^\top(X_j)) \right), \end{aligned}$$

and

$$Y_{t+L}^i = \left(\Pi_{k=t}^{t+L-1}(I + \alpha_k A(X_k))\right) \sum_{j=1}^{N} w_{t:t+L-1}^{ij} Y_t^j + \tilde{b}_{t+L}^i,$$

where

$$\begin{aligned} \tilde{b}_{t+L}^i = {} & \alpha_{t+L-1}(b^i(X_{t+L-1}) - B(X_{t+L-1})^\top \pi_{t+L}) \\ & + \sum_{k=t}^{t+L-2} \alpha_k \left(\Pi_{j=k+1}^{t+L-1} (I + \alpha_j A(X_j)) \right) \sum_{j=1}^{N} w_{k+1:t+L-1}^{ij} (b^j(X_k) - B(X_k)^\top \pi_{k+1}). \end{aligned}$$

To prove the theorem, we need the following lemmas.

Lemma 10 *Suppose that Assumptions 1 and 2 hold and $\{\mathbb{G}_t\}$ is uniformly strongly connected by sub-sequences of length L. Given α_t and T_2 defined in Theorem 3, for all $t \geq T_2 L$,*

$$\begin{aligned} & \sum_{i=1}^{N} \pi_t^i \|\theta_t^i - \langle\theta\rangle_t\|_2^2 \\ & \leq \epsilon^{q_t - T_2} \sum_{i=1}^{N} \pi_{T_2L+m_t}^i \|\theta_{T_2L+m_t}^i - \langle\theta\rangle_{T_2L+m_t}\|_2^2 + \frac{\zeta_6}{1-\epsilon}\left(\epsilon^{\frac{q_t-1}{2}} \alpha_{m_t} + \alpha_{\lceil \frac{q_t-1}{2} \rceil L + m_t}\right) \\ & \leq \epsilon^{q_t - T_2} \sum_{i=1}^{N} \pi_{T_2L+m_t}^i \|\theta_{T_2L+m_t}^i - \langle\theta\rangle_{T_2L+m_t}\|_2^2 + \frac{\zeta_6}{1-\epsilon}\left(\alpha_0 \epsilon^{\frac{q_t-1}{2}} + \alpha_{\lceil \frac{q_t-1}{2} \rceil L}\right), \end{aligned}$$

where ϵ and ζ_6 are defined in equation 6 and equation 16, respectively.

Proof of Lemma 10: Similar to the proof of Lemma 5, we have

$$\|Y_{t+L}\|_{M_{t+L}}^2 = \sum_{i=1}^{N} \pi_{t+L}^i \| \left(\Pi_{k=t}^{t+L-1}(I + \alpha_k A(X_k))\right) \sum_{j=1}^{N} w_{t:t+L-1}^{ij} Y_t^j \|_2^2 \tag{64}$$

$$+ \sum_{i=1}^{N} \pi_{t+L}^i \|\tilde{b}_{t+L}^i\|_2^2 \tag{65}$$

$$+ 2\sum_{i=1}^{N} \pi_{t+L}^i (\tilde{b}_{t+L}^i)^\top \left(\Pi_{k=t}^{t+L-1}(I + \alpha_k A(X_k))\right) \sum_{j=1}^{N} w_{t:t+L-1}^{ij} Y_t^j. \tag{66}$$

By using Lemma 4, the item given by equation 64 can be bounded as follows:

$$
\begin{aligned}
&\sum_{i=1}^{N}\pi_{t+L}^{i}\|\left(\Pi_{k=t}^{t+L-1}(I+\alpha_k A(X_k))\right)\sum_{j=1}^{N}w_{t:t+L-1}^{ij}Y_t^j\|_2^2 \\
&\leq \Pi_{k=t}^{t+L-1}(1+\alpha_k A_{\max})^2\left[\sum_{i=1}^{N}\pi_t^i\|Y_t^i\|_2^2-\frac{1}{2}\sum_{i=1}^{N}\pi_{t+L}^i\sum_{j=1}^{N}\sum_{l=1}^{N}w_{t:t+L-1}^{ij}w_{t:t+L-1}^{il}\|Y_t^j-Y_t^l\|_2^2\right] \\
&\leq \Pi_{k=t}^{t+L-1}(1+\alpha_k A_{\max})^2(1-\frac{\pi_{\min}\beta^{2L}}{2\delta_{\max}})\sum_{i=1}^{N}\pi_t^i\|Y_t^i\|_2^2. \qquad (67)
\end{aligned}
$$

Since $\|b^i(X_t)-B(X_t)^\top\pi_{t+1}\|_2\leq 2b_{\max}$ holds for all i, then

$$
\begin{aligned}
&\|\tilde{b}_{t+L}^i\|_2 \\
&\leq \alpha_{t+L-1}\|(b^i(X_{t+L-1})-B(X_{t+L-1})^\top\pi_{t+L})\|_2 \\
&\quad+\sum_{k=t}^{t+L-2}\alpha_k\|\left(\Pi_{j=k+1}^{t+L-1}(I+\alpha_j A(X_j))\right)\|_2\sum_{j=1}^{N}w_{k+1:t+L-1}^{ij}\|(b^j(X_k)-B(X_k)^\top\pi_{k+1})\|_2 \\
&\leq 2b_{\max}\left[\alpha_{t+L-1}+\sum_{k=t}^{t+L-2}\alpha_k\left(\Pi_{j=k+1}^{t+L-1}(1+\alpha_j A_{\max})\right)\right].
\end{aligned}
$$

Then, we can bound the item given by equation 65 as follows:

$$
\sum_{i=1}^{N}\pi_{t+L}^i\|\tilde{b}_{t+L}^i\|_2^2\leq 4b_{\max}^2\left(\alpha_{t+L-1}+\sum_{k=t}^{t+L-2}\alpha_k\left(\Pi_{j=k+1}^{t+L-1}(1+\alpha_j A_{\max})\right)\right)^2. \qquad (68)
$$

As for the item given by equation 66, we have

$$
\begin{aligned}
&2\sum_{i=1}^{N}\pi_{t+L}^i(\tilde{b}_{t+L}^i)^\top\left(\Pi_{k=t}^{t+L-1}(I+\alpha_k A(X_k))\right)\sum_{j=1}^{N}w_{t:t+L-1}^{ij}Y_t^j \\
&\leq 2\sum_{i=1}^{N}\pi_{t+L}^i\|\tilde{b}_{t+L}^i\|_2\|\Pi_{k=t}^{t+L-1}(I+\alpha_k A(X_k))\|_2\sum_{j=1}^{N}w_{t:t+L-1}^{ij}\|Y_t^j\|_2 \\
&\leq 2b_{\max}\left(\alpha_{t+L-1}+\sum_{k=t}^{t+L-2}\alpha_k\left(\Pi_{j=k+1}^{t+L-1}(1+\alpha_j A_{\max})\right)\right)\left(\Pi_{k=t}^{t+L-1}(I+\alpha_k A_{\max})\right)\times \\
&\left(\sum_{i=1}^{N}\pi_t^i\|Y_t^i\|_2^2+1\right). \qquad (69)
\end{aligned}
$$

From equation 67–equation 69, we have

$$
\begin{aligned}
&\|Y_{t+L}\|^2_{M_{t+L}}\\
&\leq \Pi_{k=t}^{t+L-1}(1+\alpha_k A_{\max})^2(1-\frac{\pi_{\min}\beta^{2L}}{2\delta_{\max}})\sum_{i=1}^N \pi_t^i\|Y_t^i\|_2^2\\
&\quad + 4b_{\max}^2\left(\alpha_{t+L-1}+\sum_{k=t}^{t+L-2}\alpha_k\left(\Pi_{j=k+1}^{t+L-1}(1+\alpha_j A_{\max})\right)\right)^2\\
&\quad + 2b_{\max}\left(\alpha_{t+L-1}+\sum_{k=t}^{t+L-2}\alpha_k\left(\Pi_{j=k+1}^{t+L-1}(1+\alpha_j A_{\max})\right)\right)\left(\Pi_{k=t}^{t+L-1}(I+\alpha_k A_{\max})\right)\\
&\quad \times\left(\sum_{i=1}^N \pi_t^i\|Y_t^i\|_2^2+1\right)\\
&=\Bigg(2b_{\max}\left(\alpha_{t+L-1}+\sum_{k=t}^{t+L-2}\alpha_k\left(\Pi_{j=k+1}^{t+L-1}(1+\alpha_j A_{\max})\right)\right)\left(\Pi_{k=t}^{t+L-1}(I+\alpha_k A_{\max})\right)\\
&\quad + \Pi_{k=t}^{t+L-1}(1+\alpha_k A_{\max})^2(1-\frac{\pi_{\min}\beta^{2L}}{2\delta_{\max}})\Bigg)\|Y_t\|^2_{M_t}\\
&\quad + 4b_{\max}^2\left(\alpha_{t+L-1}+\sum_{k=t}^{t+L-2}\alpha_k\left(\Pi_{j=k+1}^{t+L-1}(1+\alpha_j A_{\max})\right)\right)^2\\
&\quad + 2b_{\max}\left(\alpha_{t+L-1}+\sum_{k=t}^{t+L-2}\alpha_k\left(\Pi_{j=k+1}^{t+L-1}(1+\alpha_j A_{\max})\right)\right)\left(\Pi_{k=t}^{t+L-1}(I+\alpha_k A_{\max})\right)\\
&=\epsilon_t\|Y_t\|^2_{M_t}+4b_{\max}^2\left(\alpha_{t+L-1}+\sum_{k=t}^{t+L-2}\alpha_k\left(\Pi_{j=k+1}^{t+L-1}(1+\alpha_j A_{\max})\right)\right)^2\\
&\quad + 2b_{\max}\left(\alpha_{t+L-1}+\sum_{k=t}^{t+L-2}\alpha_k\left(\Pi_{j=k+1}^{t+L-1}(1+\alpha_j A_{\max})\right)\right)\left(\Pi_{k=t}^{t+L-1}(I+\alpha_k A_{\max})\right),
\end{aligned}
$$

where

$$
\begin{aligned}
\epsilon_t &= 2b_{\max}\left(\alpha_{t+L-1}+\sum_{k=t}^{t+L-2}\alpha_k\left(\Pi_{j=k+1}^{t+L-1}(1+\alpha_j A_{\max})\right)\right)\left(\Pi_{k=t}^{t+L-1}(I+\alpha_k A_{\max})\right)\\
&\quad + \Pi_{k=t}^{t+L-1}(1+\alpha_k A_{\max})^2(1-\frac{\pi_{\min}\beta^{2L}}{2\delta_{\max}}).
\end{aligned}
$$

Since for all $t\geq T_2L$, we have $\alpha_t\leq\alpha$, then for $t\geq T_2L$ we have $0\leq\epsilon_t\leq\epsilon\leq 1$ and

$$
\begin{aligned}
\alpha_{t+L-1}+\sum_{k=t}^{t+L-2}\alpha_k\left(\Pi_{j=k+1}^{t+L-1}(1+\alpha_j A_{\max})\right) &\leq \sum_{k=t}^{t+L-1}\alpha_k(1+\alpha A_{\max})^{t+L-k-1}\\
&\leq (1+\alpha A_{\max})^{L-1}\sum_{k=t}^{t+L-1}\alpha_k.
\end{aligned}
$$

Since we have $\sum_{k=t}^{t+L-1}\alpha_k\leq L\alpha_t\leq L\alpha$. Then, we can get

$$
\begin{aligned}
&\|Y_{t+L}\|^2_{M_{t+L}}\\
&\leq \epsilon\|Y_t\|^2_{M_t}+4b_{\max}^2(1+\alpha A_{\max})^{2L-2}\left(\sum_{k=t}^{t+L-1}\alpha_k\right)^2+2b_{\max}(1+\alpha A_{\max})^{2L-1}\left(\sum_{k=t}^{t+L-1}\alpha_k\right)\\
&\leq \epsilon\|Y_t\|^2_{M_t}+\left(4b_{\max}^2\alpha L^2(1+\alpha A_{\max})^{2L-2}+2b_{\max}L(1+\alpha A_{\max})^{2L-1}\right)\alpha_t\\
&\leq \epsilon\|Y_t\|^2_{M_t}+\zeta_6\alpha_t,
\end{aligned}
$$

where ϵ and ζ_6 are defined in equation 6 and equation 16 respectively. Then,

$$\begin{aligned}
&\|Y_{t+L}\|_{M_{t+L}}^2 \\
&\leq \epsilon\|Y_t\|_{M_t}^2 + \zeta_6\alpha_t \\
&\leq \epsilon^{q_{t+L}-T_2}\|Y_{m_t+T_2L}\|_{M_{m_t+T_2L}}^2 + \zeta_6\sum_{k=T_2}^{q_t}\epsilon^{q_t-k}\alpha_{kL+m_t} \\
&\leq \epsilon^{q_{t+L}-T_2}\|Y_{T_2L+m_t}\|_{M_{T_2L+m_t}}^2 + \zeta_6\left(\sum_{k=0}^{\lfloor\frac{q_t}{2}\rfloor}\epsilon^{q_t-k}\alpha_{kL+m_t} + \sum_{k=\lceil\frac{q_t}{2}\rceil}^{q_t}\epsilon^{q_t-k}\alpha_{kL+m_t}\right) \\
&\leq \epsilon^{q_{t+L}-T_2}\|Y_{T_2L+m_t}\|_{M_{T_2L+m_t}}^2 + \frac{\zeta_6}{1-\epsilon}\left(\epsilon^{\frac{q_t}{2}}\alpha_{m_t} + \alpha_{\lceil\frac{q_t}{2}\rceil L+m_t}\right),
\end{aligned}$$

which implies

$$\begin{aligned}
&\sum_{i=1}^N \pi_t^i\|\theta_t^i - \langle\theta\rangle_t\|_2^2 \\
&\leq \epsilon^{q_t-T_2}\sum_{i=1}^N \pi_{T_2L+m_t}^i\|\theta_{T_2L+m_t}^i - \langle\theta\rangle_{T_2L+m_t}\|_2^2 + \frac{\zeta_6}{1-\epsilon}\left(\epsilon^{\frac{q_t-1}{2}}\alpha_{m_t} + \alpha_{\lceil\frac{q_t-1}{2}\rceil L+m_t}\right) \\
&\leq \epsilon^{q_t-T_2}\sum_{i=1}^N \pi_{T_2L+m_t}^i\|\theta_{T_2L+m_t}^i - \langle\theta\rangle_{T_2L+m_t}\|_2^2 + \frac{\zeta_6}{1-\epsilon}\left(\alpha_0\epsilon^{\frac{q_t-1}{2}} + \alpha_{\lceil\frac{q_t-1}{2}\rceil L}\right).
\end{aligned}$$

This completes the proof. ∎

Lemma 11 *Suppose that Assumptions 2 and 3 hold. When the step-size α_t and corresponding mixing time $\tau(\alpha_t)$ satisfy*

$$0 < \alpha_t\tau(\alpha_t) < \frac{\log 2}{A_{\max}},$$

we have for any $t \geq T_2L$,

$$\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 \leq 2A_{\max}\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k + 2b_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k, \tag{70}$$

$$\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 \leq 6A_{\max}\|\langle\theta\rangle_t\|_2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k + 5b_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k, \tag{71}$$

$$\begin{aligned}
\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2 &\leq 72\alpha_{t-\tau(\alpha_t)}^2\tau^2(\alpha_t)A_{\max}^2\|\langle\theta\rangle_t\|_2^2 + 50\alpha_{t-\tau(\alpha_t)}^2\tau^2(\alpha_t)b_{\max}^2 \\
&\leq 8\|\langle\theta\rangle_t\|_2^2 + \frac{6b_{\max}^2}{A_{\max}^2}.
\end{aligned} \tag{72}$$

Proof of Lemma 11: Recall the update of $\langle\theta\rangle_t$ in equation 4:

$$\langle\theta\rangle_{t+1} = \langle\theta\rangle_t + \alpha_tA(X_t)\langle\theta\rangle_t + \alpha_tB(X_t)^\top\pi_{t+1}.$$

Then, we have

$$\|\langle\theta\rangle_{t+1}\|_2 \leq \|\langle\theta\rangle_t\|_2 + \alpha_tA_{\max}\|\langle\theta\rangle_t\|_2 + \alpha_tb_{\max} \leq (1+\alpha_tA_{\max})\|\langle\theta\rangle_t\|_2 + \alpha_tb_{\max}.$$

Similar to the proof of Lemma 6, for all $u \in [t-\tau(\alpha_t), t]$, we have

$$
\begin{aligned}
&\|\langle\theta\rangle_u\|_2 \\
&\leq \Pi_{k=t-\tau(\alpha_t)}^{u-1}(1+\alpha_k A_{\max})\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 + b_{\max}\sum_{k=t-\tau(\alpha_t)}^{u-1}\alpha_k \Pi_{l=k+1}^{u-1}(1+\alpha_l A_{\max}) \\
&\leq \exp\{\sum_{k=t-\tau(\alpha_t)}^{u-1}\alpha_k A_{\max}\}\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 + b_{\max}\sum_{k=t-\tau(\alpha_t)}^{u-1}\alpha_k \exp\{\sum_{l=k+1}^{u-1}\alpha_l A_{\max}\} \\
&\leq \exp\{\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)A_{\max}\}\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 + b_{\max}\sum_{k=t-\tau(\alpha_t)}^{u-1}\alpha_k \exp\{\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)A_{\max}\} \\
&\leq 2\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 + 2b_{\max}\sum_{k=t-\tau(\alpha_t)}^{u-1}\alpha_k,
\end{aligned}
$$

where we use $\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)A_{\max} \leq \log 2 < \frac{1}{3}$ in the last inequality. Thus, for all $t \geq T_2 L$, we can get equation 70 as follows:

$$
\begin{aligned}
&\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 \\
&\leq \sum_{k=t-\tau(\alpha_t)}^{t-1}\|\langle\theta\rangle_{k+1} - \langle\theta\rangle_k\|_2 \\
&\leq A_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k\|\langle\theta\rangle_k\|_2 + b_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\leq A_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k\left(2\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 + 2b_{\max}\sum_{l=t-\tau(\alpha_t)}^{k-1}\alpha_l\right) + b_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\leq 2A_{\max}\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k + \left(2A_{\max}\tau(\alpha_t)\alpha_{t-\tau(\alpha_t)}+1\right)b_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\leq 2A_{\max}\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k + \frac{5}{3}b_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\leq 2A_{\max}\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k + 2b_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k.
\end{aligned}
$$

Moreover, by using the above inequality, we can get equation 71 for all $t \geq T_2 L$ as follows:

$$
\begin{aligned}
&\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 \\
&\leq 2A_{\max}\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k + \frac{5}{3}b_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\leq 2A_{\max}\tau(\alpha_t)\alpha_{t-\tau(\alpha_t)}\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 + 2A_{\max}\|\langle\theta\rangle_t\|_2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\quad + \frac{5}{3}b_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\leq 6A_{\max}\|\langle\theta\rangle_t\|_2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k + 5b_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k.
\end{aligned}
$$

Next, by using equation 71 and the inequality $(x+y)^2 \leq 2x^2+y^2$ for all x, y, we can get equation 72 as follows:

$$\begin{aligned}
\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2 &\leq 72A_{\max}^2\|\langle\theta\rangle_t\|_2^2(\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k)^2 + 50b_{\max}^2(\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k)^2 \\
&\leq 72\alpha_{t-\tau(\alpha_t)}^2\tau^2(\alpha_t)A_{\max}^2\|\langle\theta\rangle_t\|_2^2 + 50\alpha_{t-\tau(\alpha_t)}^2\tau^2(\alpha_t)b_{\max}^2 \\
&\leq 8\|\langle\theta\rangle_t\|_2^2 + 6(\frac{b_{\max}}{A_{\max}})^2,
\end{aligned}$$

where we use $\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)A_{\max} < \frac{1}{3}$ in the last inequality. ■

Lemma 12 *Suppose that Assumptions 2–6 hold and $\{\mathbb{G}_t\}$ is uniformly strongly connected. When*

$$0 < \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t) < \frac{\log 2}{A_{\max}},$$

we have for any $t \geq T_2 L$,

$$\begin{aligned}
&\left|\mathbf{E}\left[(\langle\theta\rangle_t - \theta^*)^\top(P+P^\top)\left(A(X_t)\langle\theta\rangle_t + B(X_t)^\top\pi_{t+1} - A\langle\theta\rangle_t - b\right) \mid \mathcal{F}_{t-\tau(\alpha_t)}\right]\right| \\
&\leq \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\left(72 + 456A_{\max}^2 + 84A_{\max}b_{\max}\right)\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\quad + \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\Bigg[2 + 4\|\theta^*\|_2^2 + \frac{48b_{\max}^2}{A_{\max}^2} + 152\left(b_{\max} + A_{\max}\|\theta^*\|_2\right)^2 + 12A_{\max}b_{\max} \\
&\quad + 48A_{\max}b_{\max}\left(\frac{b_{\max}}{A_{\max}} + 1\right)^2 + 87b_{\max}^2\Bigg] \\
&\quad + 2\gamma_{\max}\eta_{t+1}\sqrt{N}b_{\max}\left(1 + 9\mathbf{E}\left[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha)}\right] + \frac{6b_{\max}^2}{A_{\max}^2} + \|\theta^*\|_2^2\right).
\end{aligned}$$

Proof of Lemma 12: Note that for all $t \geq T_2 L$, we have

$$\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_t - \theta^*)^\top(P+P^\top)(A(X_t)\langle\theta\rangle_t + B(X_t)^\top\pi_{t+1} - A\langle\theta\rangle_t - b) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq |\mathbf{E}[(\langle\theta\rangle_t - \theta^*)^\top(P+P^\top)(A(X_t) - A)\langle\theta\rangle_t \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\quad + |\mathbf{E}[(\langle\theta\rangle_t - \theta^*)^\top(P+P^\top)(B(X_t)^\top\pi_{t+1} - b) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq |\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top(P+P^\top)(A(X_t) - A)\langle\theta\rangle_{t-\tau(\alpha_t)} \mid \mathcal{F}_{t-\tau(\alpha_t)}]| && (73)\\
&\quad + |\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top(P+P^\top)(A(X_t) - A)(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| && (74)\\
&\quad + |\mathbf{E}[(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)})^\top(P+P^\top)(A(X_t) - A)\langle\theta\rangle_{t-\tau(\alpha_t)} \mid \mathcal{F}_{t-\tau(\alpha_t)}]| && (75)\\
&\quad + |\mathbf{E}[(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)})^\top(P+P^\top)(A(X_t) - A)(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| && (76)\\
&\quad + |\mathbf{E}[(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)})^\top(P+P^\top)(B(X_t)^\top\pi_{t+1} - b) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| && (77)\\
&\quad + |\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top(P+P^\top)(B(X_t)^\top\pi_{t+1} - b) \mid \mathcal{F}_{t-\tau(\alpha_t)}]|. && (78)
\end{aligned}$$

Similar to the proof of Lemma 7, by using the mixing time in Assumption 3, we can get the bound for equation 73 and equation 78 for all $t \geq T_2 L$:

$$\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top(P+P^\top)(A(X_t) - A)\langle\theta\rangle_{t-\tau(\alpha_t)} \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq |(\langle\theta\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top(P+P^\top)\mathbf{E}[A(X_t) - A \mid \mathcal{F}_{t-\tau(\alpha_t)}]\langle\theta\rangle_{t-\tau(\alpha_t)}| \\
&\leq 2\alpha_t\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha_t)} - \theta^*\|_2\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq \alpha_t\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha_t)} - \theta^*\|_2^2 + \|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq \alpha_t\gamma_{\max}\mathbf{E}[2\|\theta^*\|_2^2 + 3\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq 6\alpha_t\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] + 6\alpha_t\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] + 2\alpha_t\gamma_{\max}\|\theta^*\|_2^2 \\
&\leq 54\alpha_t\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] + 36\alpha_t\gamma_{\max}(\frac{b_{\max}}{A_{\max}})^2 + 2\alpha_t\gamma_{\max}\|\theta^*\|_2^2, && (79)
\end{aligned}$$

where in the last inequality, we use equation 72 from Lemma 11.

$$
\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha_t)}-\theta^*)^\top(P+P^\top)(B(X_t)^\top\pi_{t+1}-b)\mid\mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\le |\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha_t)}-\theta^*)^\top(P+P^\top)(\sum_{i=1}^N\pi_{t+1}^i(b^i(X_t)-b^i)+\sum_{i=1}^N(\pi_{t+1}^i-\pi_\infty^i)b^i)\mid\mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\le |(\langle\theta\rangle_{t-\tau(\alpha_t)}-\theta^*)^\top(P+P^\top)(\sum_{i=1}^N\pi_{t+1}^i\mathbf{E}[b^i(X_t)-b^i\mid\mathcal{F}_{t-\tau(\alpha_t)}]+\sum_{i=1}^N(\pi_{t+1}^i-\pi_\infty^i)b^i)| \\
&\le 2\gamma_{\max}(\alpha_t+\eta_{t+1}\sqrt{N}b_{\max})\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha_t)}-\theta^*\|_2\mid\mathcal{F}_{t-\tau(\alpha_t)}] \\
&\le 2\gamma_{\max}(\alpha_t+\eta_{t+1}\sqrt{N}b_{\max})\left(1+\frac{1}{2}\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]+\frac{1}{2}\|\theta^*\|_2^2\right) \\
&\le 2\gamma_{\max}(\alpha_t+\eta_{t+1}\sqrt{N}b_{\max})\left(1+\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha_t)}-\langle\theta\rangle_t\|_2^2+\|\langle\theta\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]+\|\theta^*\|_2^2\right) \\
&\le 2\gamma_{\max}(\alpha_t+\eta_{t+1}\sqrt{N}b_{\max})\left(1+9\mathbf{E}[\|\langle\theta\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]+6(\frac{b_{\max}}{A_{\max}})^2+\|\theta^*\|_2^2\right),
\end{aligned}
\tag{80}
$$

where in the last inequality we use equation 72.

Next, by using Assumption 2, equation 70 and equation 72, we have

$$
\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_{t-\tau(\alpha_t)}-\theta^*)^\top(P+P^\top)(A(X_t)-A)(\langle\theta\rangle_t-\langle\theta\rangle_{t-\tau(\alpha_t)})\mid\mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\le 4\gamma_{\max}A_{\max}\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha_t)}-\theta^*\|_2\|\langle\theta\rangle_t-\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2\mid\mathcal{F}_{t-\tau(\alpha_t)}] \\
&\le 4\gamma_{\max}A_{\max}\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2\|\langle\theta\rangle_t-\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2+\|\theta^*\|_2\|\langle\theta\rangle_t-\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2\mid\mathcal{F}_{t-\tau(\alpha_t)}] \\
&\le 8\gamma_{\max}A_{\max}^2\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k+8\gamma_{\max}A_{\max}b_{\max}\|\theta^*\|_2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\quad+8\gamma_{\max}A_{\max}^2\left(\frac{b_{\max}}{A_{\max}}+\|\theta^*\|_2\right)\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2\mid\mathcal{F}_{t-\tau(\alpha_t)}]\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\le \gamma_{\max}A_{\max}^2\left(12\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]+8\left(\frac{b_{\max}}{A_{\max}}+\|\theta^*\|_2\right)^2\right)\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\le 24\gamma_{\max}A_{\max}^2\mathbf{E}[\|\langle\theta\rangle_t-\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\quad+24\gamma_{\max}A_{\max}^2\mathbf{E}[\|\langle\theta\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\quad+8\gamma_{\max}A_{\max}^2\left(\frac{b_{\max}}{A_{\max}}+\|\theta^*\|_2\right)^2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\le \gamma_{\max}\left(216A_{\max}^2\mathbf{E}[\|\langle\theta\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]+152\left(b_{\max}+A_{\max}\|\theta^*\|_2\right)^2\right)\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k.
\end{aligned}
\tag{81}
$$

In additional, as for the bound of equation 75, by using equation 70 and equation 72, we have

$$
\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)})^\top (P+P^\top)(A(X_t)-A)\langle\theta\rangle_{t-\tau(\alpha_t)} \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq 4\gamma_{\max} A_{\max} \mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 \|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq 8\gamma_{\max} A_{\max} \mathbf{E}[A_{\max}\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2 + b_{\max}\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\leq 4\gamma_{\max} A_{\max} \left((2A_{\max}+b_{\max})\mathbf{E}[\|\langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] + b_{\max}\right) \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\leq 8\gamma_{\max} A_{\max}(2A_{\max}+b_{\max})\mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 8\gamma_{\max} A_{\max}(2A_{\max}+b_{\max})\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 4\gamma_{\max} A_{\max} b_{\max} \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\leq 72\, \gamma_{\max} A_{\max}(2A_{\max}+b_{\max})\mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 48\gamma_{\max} A_{\max} b_{\max} (\frac{b_{\max}}{A_{\max}}+1)^2 \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k.
\end{aligned} \tag{82}
$$

Moreover, by using equation 72, we can get the bound for equation 76 as follows:

$$
\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)})^\top (P+P^\top)(A(X_t)-A)(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq 4\gamma_{\max} A_{\max} \mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq 4\gamma_{\max} A_{\max} \mathbf{E}[72A_{\max}^2\|\langle\theta\rangle_t\|_2^2 + 50b_{\max}^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}](\sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k)^2 \\
&\leq 96A_{\max}^2 \gamma_{\max} \mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k + 67b_{\max}^2 \gamma_{\max} \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k.
\end{aligned} \tag{83}
$$

Finally, we can get the bound of equation 77 by using equation 71:

$$
\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)})(P+P^\top)(B(X_t)^\top \pi_{t+1} - b) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq 4\gamma_{\max} b_{\max} \mathbf{E}[\|\langle\theta\rangle_t - \langle\theta\rangle_{t-\tau(\alpha_t)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq 4\gamma_{\max} b_{\max} \mathbf{E}[6A_{\max}\|\langle\theta\rangle_t\|_2 + 5b_{\max} \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\leq \gamma_{\max} \left(12A_{\max} b_{\max} \mathbf{E}[\|\langle\theta\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] + 12A_{\max} b_{\max} + 20b_{\max}^2\right) \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k.
\end{aligned} \tag{84}
$$

Then, by using equation 79–equation 84, we have

$$
\begin{aligned}
&|\mathbf{E}[(\langle\theta\rangle_t-\theta^*)^\top(P+P^\top)(A(X_t)\langle\theta\rangle_t+B(X_t)^\top\pi_{t+1}-A\langle\theta\rangle_t-b)\mid\mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq 54\alpha_t\gamma_{\max}\mathbf{E}[\|\langle\theta\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]+36\alpha_t\gamma_{\max}(\frac{b_{\max}}{A_{\max}})^2+2\alpha_t\gamma_{\max}\|\theta^*\|_2^2 \\
&\quad+216\gamma_{\max}A_{\max}^2\mathbf{E}[\|\langle\theta\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\quad+152\gamma_{\max}\left(b_{\max}+A_{\max}\|\theta^*\|_2\right)^2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k+67b_{\max}^2\gamma_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\quad+12\,\gamma_{\max}A_{\max}(20A_{\max}+7b_{\max})\mathbf{E}[\|\langle\theta\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\quad+48\gamma_{\max}A_{\max}b_{\max}(\frac{b_{\max}}{A_{\max}}+1)^2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k+(12A_{\max}b_{\max}+20b_{\max}^2)\gamma_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k \\
&\quad+2\gamma_{\max}(\alpha_t+\eta_{t+1}\sqrt{N}b_{\max})\left(1+9\mathbf{E}[\|\langle\theta\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]+6(\frac{b_{\max}}{A_{\max}})^2+\|\theta^*\|_2^2\right) \\
&\leq\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\left(72+456A_{\max}^2+84A_{\max}b_{\max}\right)\mathbf{E}[\|\langle\theta\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}] \\
&\quad+\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\Bigg[2+4\|\theta^*\|_2^2+48(\frac{b_{\max}}{A_{\max}})^2+152\left(b_{\max}+A_{\max}\|\theta^*\|_2\right)^2+12A_{\max}b_{\max} \\
&\quad+48A_{\max}b_{\max}(\frac{b_{\max}}{A_{\max}}+1)^2+87b_{\max}^2\Bigg] \\
&\quad+2\gamma_{\max}\eta_{t+1}\sqrt{N}b_{\max}\left(1+9\mathbf{E}[\|\langle\theta\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha)}]+6(\frac{b_{\max}}{A_{\max}})^2+\|\theta^*\|_2^2\right),
\end{aligned}
$$

where we use $\alpha_t\leq\alpha_{t-\tau(\alpha_t)}$ from Assumption 5 and $\tau(\alpha_t)\geq 1$ in the last inequality. This completes the proof. ■

Lemma 13 *Under Assumptions 1–6, when the* $\tau(\alpha_t)\alpha_{t-\tau(\alpha_t)}\leq\min\{\frac{\log 2}{A_{\max}},\ \frac{0.1}{\zeta_5\gamma_{\max}}\}$*, we have for any* $t\geq T_2L$*,*

$$
\begin{aligned}
\mathbf{E}\left[\|\langle\theta\rangle_t-\theta^*\|_2^2\right]\leq\frac{T_2L}{t}\frac{\gamma_{\max}}{\gamma_{\min}}\mathbf{E}[\|\langle\theta\rangle_{T_2L}-\theta^*\|_2^2]+\frac{\zeta_7\alpha_0C\log^2(\frac{t}{\alpha_0})}{t}\frac{\gamma_{\max}}{\gamma_{\min}} \\
+\alpha_0\zeta_4\frac{\gamma_{\max}}{\gamma_{\min}}\frac{\sum_{l=T_2L}^t\eta_l}{t},
\end{aligned}
$$

where T_2 *is defined in Appendix A.1, and* ζ_4*,* ζ_5*,* ζ_7 *are defined in equation 14, equation 15, equation 17, respectively.*

Proof of Lemma 13: Recall the update of $\langle\theta\rangle_t$ in equation 4:

$$\langle\theta\rangle_{t+1}=\langle\theta\rangle_t+\alpha_tA(X_t)\langle\theta\rangle_t+\alpha_tB(X_t)^\top\pi_{t+1}.$$

Note that $\mathbf{E}[\|\langle\theta\rangle_t\|_2^2] \leq 2\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] + 2\|\theta^*\|_2^2 \leq \frac{2}{\gamma_{\min}}\mathbf{E}[H(\langle\theta\rangle_t)] + 2\|\theta^*\|_2^2$, then from equation 60 and Lemma 12, for $t \geq T_2 L$ we have

$$
\begin{aligned}
&\mathbf{E}[H(\langle\theta\rangle_{t+1})] \\
&\leq \mathbf{E}[H(\langle\theta\rangle_t)] - \alpha_t \mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] + \alpha_t^2 A_{\max}^2 \gamma_{\max} \mathbf{E}[\|\langle\theta\rangle_t\|_2^2] + \alpha_t^2 b_{\max}^2 \gamma_{\max} \\
&\quad + 2\alpha_t^2 A_{\max} b_{\max} \gamma_{\max} \mathbf{E}[\|\langle\theta\rangle_t\|_2] \\
&\quad + \alpha_t \mathbf{E}[(\langle\theta\rangle_t - \theta^*)^\top (P + P^\top)(A(X_t)\langle\theta\rangle_t + B(X_t)^\top \pi_{t+1} - A\langle\theta\rangle_t - b)] \\
&\leq \mathbf{E}[H(\langle\theta\rangle_t)] - \alpha_t \mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] + 2\alpha_t^2 A_{\max}^2 \gamma_{\max} \mathbf{E}[\|\langle\theta\rangle_t\|_2^2] + 2\alpha_t^2 b_{\max}^2 \gamma_{\max} \\
&\quad + \alpha_t \alpha_{t-\tau(\alpha_t)} \tau(\alpha_t) \gamma_{\max} \left(72 + 456 A_{\max}^2 + 84 A_{\max} b_{\max}\right) \mathbf{E}[\|\langle\theta\rangle_t\|_2^2] \\
&\quad + \alpha_t \alpha_{t-\tau(\alpha_t)} \tau(\alpha_t) \gamma_{\max} \Bigg[2 + 4\|\theta^*\|_2^2 + 48(\frac{b_{\max}}{A_{\max}})^2 + 152\left(b_{\max} + A_{\max}\|\theta^*\|_2\right)^2 \\
&\quad + 12 A_{\max} b_{\max} + 48 A_{\max} b_{\max} (\frac{b_{\max}}{A_{\max}} + 1)^2 + 87 b_{\max}^2\Bigg] \\
&\quad + 2\alpha_t \gamma_{\max} \eta_{t+1} \sqrt{N} b_{\max} \left(1 + 9\mathbf{E}[\|\langle\theta\rangle_t\|_2^2] + 6(\frac{b_{\max}}{A_{\max}})^2 + \|\theta^*\|_2^2\right) \\
&\leq \mathbf{E}[H(\langle\theta\rangle_t)] + 2\alpha_t \alpha_{t-\tau(\alpha_t)} \tau(\alpha_t) \gamma_{\max} \left(72 + 458 A_{\max}^2 + 84 A_{\max} b_{\max}\right) \|\theta^*\|_2^2 \\
&\quad + \left(-\alpha_t + 2\alpha_t \alpha_{t-\tau(\alpha_t)} \tau(\alpha_t) \gamma_{\max} \left(72 + 458 A_{\max}^2 + 84 A_{\max} b_{\max}\right)\right) \mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] \\
&\quad + \alpha_t \alpha_{t-\tau(\alpha_t)} \tau(\alpha_t) \gamma_{\max} \Bigg[2 + 4\|\theta^*\|_2^2 + 48(\frac{b_{\max}}{A_{\max}})^2 + 152\left(b_{\max} + A_{\max}\|\theta^*\|_2\right)^2 \\
&\quad + 12 A_{\max} b_{\max} + 48 A_{\max} b_{\max} (\frac{b_{\max}}{A_{\max}} + 1)^2 + 89 b_{\max}^2\Bigg] \\
&\quad + 2\alpha_t \gamma_{\max} \eta_{t+1} \sqrt{N} b_{\max} \left(1 + 18\mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] + 6(\frac{b_{\max}}{A_{\max}})^2 + 19\|\theta^*\|_2^2\right) \\
&\leq \mathbf{E}[H(\langle\theta\rangle_t)] + \left(-\alpha_t + \alpha_t \alpha_{t-\tau(\alpha_t)} \tau(\alpha_t) \gamma_{\max} \zeta_5 + 36 \alpha_t \gamma_{\max} \eta_{t+1} \sqrt{N} b_{\max}\right) \mathbf{E}[\|\langle\theta\rangle_t - \theta^*\|_2^2] \\
&\quad + \alpha_t \alpha_{t-\tau(\alpha_t)} \tau(\alpha_t) \gamma_{\max} \zeta_7 + \alpha_t \gamma_{\max} \eta_{t+1} \zeta_4,
\end{aligned}
$$

where ζ_4, ζ_5 and ζ_7 are defined in equation 14, equation 15 and equation 17, respectively. Moreover, from $\alpha_t = \frac{\alpha_0}{t+1}$, $\alpha_0 \geq \frac{\gamma_{max}}{0.9}$ and the definition of T_2, we have for all $t \geq T_2 L$

$$
\begin{aligned}
\mathbf{E}[H(\langle\theta\rangle_{t+1})] &\leq \left(1 - \frac{0.9\alpha_t}{\gamma_{\max}}\right) \mathbf{E}[H(\langle\theta\rangle_t)] + \alpha_t \gamma_{\max} \eta_{t+1} \zeta_4 + \alpha_t \alpha_{t-\tau(\alpha_t)} \tau(\alpha_t) \gamma_{\max} \zeta_7 \\
&\leq \frac{t}{t+1} \mathbf{E}[H(\langle\theta\rangle_t)] + \alpha_0 \gamma_{\max} \zeta_4 \frac{\eta_{t+1}}{t+1} + \frac{\alpha_0^2 C \log(\frac{t+1}{\alpha_0}) \gamma_{\max} \zeta_7}{(t+1)(t - \tau(\alpha_t) + 1)} \\
&\leq \frac{T_2 L}{t+1} \mathbf{E}[H(\langle\theta\rangle_{T_2 L})] + \alpha_0 \gamma_{\max} \zeta_4 \sum_{l=T_2 L}^{t} \frac{\eta_{l+1}}{l+1} \Pi_{u=l+1}^{t} \frac{u}{u+1} \\
&\quad + \alpha_0^2 \gamma_{\max} \zeta_7 \sum_{l=T_2 L}^{t} \frac{C \log(\frac{l+1}{\alpha_0})}{(l+1)(l - \tau(\alpha_l) + 1)} \Pi_{u=l+1}^{t} \frac{u}{u+1} \\
&\leq \frac{T_2 L}{t+1} \mathbf{E}[H(\langle\theta\rangle_{T_2 L})] + \alpha_0 \gamma_{\max} \zeta_4 \frac{\sum_{l=T_2 L}^{t} \eta_{l+1}}{t+1} + \frac{\zeta_7 \alpha_0 \gamma_{\max} C \log^2(\frac{t+1}{\alpha_0})}{t+1} \\
&\leq \frac{T_2 L}{t+1} \mathbf{E}[H(\langle\theta\rangle_{T_2 L})] + \alpha_0 \gamma_{\max} \zeta_4 \frac{\sum_{l=T_2 L}^{t+1} \eta_l}{t+1} + \frac{\zeta_7 \alpha_0 \gamma_{\max} C \log^2(\frac{t+1}{\alpha_0})}{t+1},
\end{aligned}
$$

where we use

$$
\sum_{l=T_2}^{t} \frac{2\alpha_0 \log(\frac{l+1}{\alpha_0})}{l+1} \leq \log^2(\frac{t+1}{\alpha_0})
$$

to get the last inequality. Then, we can get the bound of $\mathbf{E}[\|\langle\theta\rangle_{t+1}-\theta^*\|_2^2]$ as follows

$$\begin{aligned}\mathbf{E}[\|\langle\theta\rangle_{t+1}-\theta^*\|_2^2] &\le \frac{1}{\gamma_{\min}}\mathbf{E}[H(\langle\theta\rangle_{t+1})]\\ &\le \frac{T_2L}{t+1}\frac{\gamma_{\max}}{\gamma_{\min}}\mathbf{E}[\|\langle\theta\rangle_{T_2L}-\theta^*\|_2^2]+\frac{\zeta_7\alpha_0 C\log^2(\frac{t+1}{\alpha_0})}{t+1}\frac{\gamma_{\max}}{\gamma_{\min}}\\ &\quad+\alpha_0\zeta_4\frac{\gamma_{\max}}{\gamma_{\min}}\frac{\sum_{l=T_2L}^{t+1}\eta_l}{t+1}.\end{aligned}$$

This completes the proof. ■

We are now in a position to prove the time-varying step-size case in Theorem 3.

Proof of Case 2) in Theorem 3: From Lemmas 10 and 13, for any $t\ge T_2L$, we have

$$\begin{aligned}\sum_{i=1}^N\pi_t^i\mathbf{E}[\|\theta_t^i-\theta^*\|_2^2] &\le 2\sum_{i=1}^N\pi_t^i\mathbf{E}[\|\theta_t^i-\langle\theta\rangle_t\|_2^2]+2\mathbf{E}[\|\langle\theta\rangle_t-\theta^*\|_2^2]\\ &\le 2\epsilon^{q_t-T_2}\sum_{i=1}^N\pi_{T_2L+m_t}^i\mathbf{E}[\|\theta_{T_2L+m_t}^i-\langle\theta\rangle_{T_2L+m_t}\|_2^2]\\ &\quad+\frac{2T_2L}{t}\frac{\gamma_{\max}}{\gamma_{\min}}\mathbf{E}[\|\langle\theta\rangle_{T_2L}-\theta^*\|_2^2]+\frac{2\zeta_7\alpha_0C\log^2(\frac{t}{\alpha_0})}{t}\frac{\gamma_{\max}}{\gamma_{\min}}\\ &\quad+2\alpha_0\zeta_4\frac{\gamma_{\max}}{\gamma_{\min}}\frac{\sum_{l=T_2L}^t\eta_l}{t}+\frac{2\zeta_6}{1-\epsilon}(\alpha_0\epsilon^{\frac{q_t-1}{2}}+\alpha_{\lceil\frac{q_t-1}{2}\rceil L})\\ &\le 2\epsilon^{q_t-T_2}\sum_{i=1}^N\pi_{LT_2+m_t}^i\mathbf{E}\left[\left\|\theta_{LT_2+m_t}^i-\langle\theta\rangle_{LT_2+m_t}\right\|_2^2\right]\\ &\quad+C_3\left(\alpha_0\epsilon^{\frac{q_t-1}{2}}+\alpha_{\lceil\frac{q_t-1}{2}\rceil L}\right)+\frac{1}{t}\left(C_4\log^2\left(\frac{t}{\alpha_0}\right)+C_5\sum_{k=LT_2}^t\eta_k+C_6\right),\end{aligned}$$

where C_3-C_6 are defined in Appendix A.1. This completes the proof. ■

E.3 Push-SA

In this subsection, we analyze the push-based distributed stochastic approximation algorithm equation 9 and provide the proofs of the results in Section 3. We begin with the proof of asymptotic performance.

Proof of Theorem 4: From Lemma 20, since $\bar\epsilon\in(0,1)$ and $\alpha_t=\frac{\alpha_0}{t}$, we have $\lim_{t\to\infty}\|\theta_{t+1}^i-\langle\tilde\theta\rangle_t\|_2=0$, which implies that all θ_{t+1}^i, $i\in\mathcal{V}$, will reach a consensus with $\langle\tilde\theta\rangle_t$. The update of $\langle\tilde\theta\rangle_t$ is given in equation 90, which can be treated as a single-agent linear stochastic approximation whose corresponding ODE is equation 10. In addition, from Theorem 5 and Lemma 21, $\lim_{\to\infty}\sum_{i=1}^N\mathbf{E}[\|\theta_{t+1}^i-\theta^*\|_2^2]=0$, it follows that θ_{t+1}^i will converge to θ^* in mean square for all $i\in\mathcal{V}$. ■

We now analyze the finite-time performance of equation 9.

Let $\hat{W}_t$ be the matrix whose ij-th entry is $\hat{w}_t^{ij}$. Then, from equation 9 we have

$$\begin{aligned}\theta_{t+1}^i = \frac{\tilde{\theta}_{t+1}^i}{y_{t+1}^i} &= \frac{\sum_{j=1}^N \hat{w}_t^{ij}(\tilde{\theta}_t^j + \alpha_t A(X_t)\theta_t^j + \alpha_t b^j(X_t))}{y_{t+1}^i} \\ &= \sum_{j=1}^N \frac{\hat{w}_t^{ij} y_t^j}{\sum_{k=1}^N \hat{w}_t^{ik} y_t^k}\left[\frac{\tilde{\theta}_t^j}{y_t^j} + \alpha_t A(X_t)\frac{\theta_t^j}{y_t^j} + \alpha_t \frac{b^j(X_t)}{y_t^j}\right] \\ &= \sum_{j=1}^N \tilde{w}_t^{ij}\left[\theta_t^j + \alpha_t A(X_t)\frac{\theta_t^j}{y_t^j} + \alpha_t \frac{b^j(X_t)}{y_t^j}\right], \end{aligned} \tag{85}$$

where $\tilde{w}_t^{ij} = \frac{\hat{w}_t^{ij} y_t^j}{\sum_{k=1}^N \hat{w}_t^{ik} y_t^k}$ and $\tilde{W}_t = [\tilde{w}_t^{ij}]$ is a row stochastic matrix, i.e.,

$$\sum_{j=1}^N \tilde{w}_t^{ij} = \frac{\sum_{j=1}^N \hat{w}_t^{ij} y_t^j}{\sum_{k=1}^N \hat{w}_t^{ik} y_t^k} = 1, \quad \forall i.$$

Let $\Theta_t = [\theta_t^1, \cdots, \theta_t^N]^\top$ and $\tilde{\Theta}_t = [\tilde{\theta}_t^1, \cdots, \tilde{\theta}_t^N]^\top$. Then equation 9 and equation 85 can be written as

$$\tilde{\Theta}_{t+1} = \hat{W}_t\left[\tilde{\Theta}_t + \alpha_t \begin{bmatrix} (\tilde{\theta}_t^1)^\top / y_t^1 \\ \cdots \\ (\tilde{\theta}_t^N)^\top / y_t^N \end{bmatrix} A(X_t)^\top + \alpha_t B(X_t)\right] \tag{86}$$

$$\Theta_{t+1} = \tilde{W}_t\left[\Theta_t + \alpha_t \begin{bmatrix} (\theta_t^1)^\top / y_t^1 \\ \cdots \\ (\theta_t^N)^\top / y_t^N \end{bmatrix} A(X_t)^\top + \alpha_t \begin{bmatrix} (b^1(X_t))^\top / y_t^1 \\ \cdots \\ (b^N(X_t))^\top / y_t^N \end{bmatrix}\right]. \tag{87}$$

Since each matrix $\tilde{W}_t = [\tilde{w}_t^{ij}]$ is stochastic, from Lemma 1, there exists a unique absolute probability sequence $\{\tilde{\pi}_t\}$ for the matrix sequence $\{\tilde{W}_t\}$ such that $\tilde{\pi}_t^i \geq \tilde{\pi}_{\min}$ for all $i \in \mathcal{V}$ and $t \geq 0$, with the constant $\tilde{\pi}_{\min} \in (0, 1)$.

Lemma 14 *Suppose that $\{\mathbb{G}_t\}$ is uniformly strongly connected. Then, $\Pi_{s=0}^t \hat{W}_s$ will converge to the set $\{v\mathbf{1}_N^\top \,:\, v \in \mathbb{R}^N\}$ exponentially fast as $t \to \infty$.*

Proof of Lemma 14: The lemma is a direct consequence of Theorem 2 in Hajnal & Bartlett (1958). ■

Lemma 15 *Suppose that $\{\mathbb{G}_t\}$ is uniformly strongly connected. Then, $(\Pi_{l=s}^t \tilde{W}_l)^{ij} = \frac{y_s^j}{y_{t+1}^i}(\Pi_{l=s}^t \hat{W}_l)^{ij}$ and $\frac{\tilde{\pi}_s^i}{y_s^i} = \frac{1}{y_s^i}\lim_{t\to\infty}(\Pi_{l=s}^t \tilde{W}_l)^{ji} = \frac{1}{N}$ for all $i, j \in \mathcal{V}$ and $s \geq 0$.*

Proof of Lemma 15: Note that for all $l \geq 0$, we have $\tilde{w}_l^{ij} = \frac{\hat{w}_l^{ij} y_l^j}{y_{l+1}^i}$. Let $\hat{W}_{s:t} = \Pi_{l=s}^t \hat{W}_l$ for all $t \geq s \geq 0$. We claim that

$$(\Pi_{l=s}^t \tilde{W}_l)^{ij} = \frac{y_s^j \hat{w}_{s:t}^{ij}}{y_{t+1}^i},$$

where $\hat{w}_{s:t}^{ij}$ is the i, j-th entry of the matrix $\hat{W}_{s:t}^{ij}$. The claim will be proved by induction on t. When $t = s + 1$,

$$\begin{aligned}(\tilde{W}_{s+1}\tilde{W}_s)^{ij} &= \sum_{k=1}^N \tilde{w}_{s+1}^{ik} \cdot \tilde{w}_s^{kj} \\ &= \sum_{k=1}^N \frac{y_{s+1}^k \hat{w}_{s+1}^{ik}}{y_{s+2}^i} \frac{y_s^j \hat{w}_s^{kj}}{y_{s+1}^k} \\ &= \frac{y_s^j}{y_{s+2}^i} \sum_{k=1}^N \hat{w}_{s+1}^{ik} \hat{w}_s^{kj} = \frac{y_s^j}{y_{s+2}^i} \hat{w}_{s:s+1}^{ij}. \end{aligned}$$

Thus, in this case the claim is true. Now suppose that the claim holds for all $t = \tau \geq s$, where τ is a positive integer. For $t = \tau + 1$, we have

$$\begin{aligned}(\Pi_{s=1}^{\tau+1}\tilde{W}_s)^{ij} &= \sum_{k=1}^{N} \tilde{w}_{\tau+1}^{ik} \cdot \frac{y_s^j \hat{w}_{s:\tau}^{kj}}{y_{\tau+1}^k} \\ &= \sum_{k=1}^{N} \frac{\hat{w}_{\tau+1}^{ik} y_{\tau+1}^k}{y_{\tau+2}^i} \cdot \frac{y_s^j \hat{w}_{s:\tau}^{kj}}{y_{\tau+1}^k} \\ &= \frac{y_s^j}{y_{\tau+2}^i} \sum_{k=1}^{N} \hat{w}_{\tau+1}^{ik} \cdot \hat{w}_{s:\tau}^{kj} = \frac{y_s^j}{y_{\tau+2}^i} \hat{w}_{s:\tau+1}^{ij},\end{aligned}$$

which establishes the claim by induction.

From Lemma 14, for given $s \geq 0$, we have $\lim_{t\to\infty} \hat{W}_{s:t} = v_{s,\infty}\mathbf{1}_N^\top$, with the understanding here that $v_{s,\infty}$ is not a constant vector. Then, since $y_{t+1} = \hat{W}_t y_t = \Pi_{l=s}^t \hat{W}_l y_s$ for all $t \geq s$, we have

$$\begin{aligned}\lim_{t\to\infty} (\Pi_{l=s}^t \tilde{W}_l)^{ij} &= \lim_{t\to\infty} \frac{y_s^j \hat{w}_{s:t}^{ij}}{y_{t+1}^i} = \lim_{t\to\infty} \frac{y_s^j \hat{w}_{s:t}^{ij}}{\sum_{k=1}^N \hat{W}_{s:t}^{ik} y_s^k} = \frac{y_s^j \lim_{t\to\infty} \hat{w}_{s:t}^{ij}}{\lim_{t\to\infty}\sum_{k=1}^N \hat{W}_{s:t}^{ik} y_s^k} \\ &= \frac{y_s^j v_{s,\infty}^i}{\sum_{k=1}^N v_{s,\infty}^i y_s^k} = \frac{y_s^j}{N},\end{aligned}$$

where we use the fact that $\mathbf{1}_N^\top y_s = N$ for all $s \geq 0$ in the last equality. This completes the proof. ■

To proceed, let

$$\begin{aligned}h^j(\Theta_n, y_n) &= A\frac{\theta_n^i}{y_n^i} + \frac{b^i}{y_n^i} \\ M_n^j &= \left(A(X_n) - \mathbb{E}[A(X_n)|\mathcal{F}_{n-\tau(\alpha_n)}]\right)\frac{\theta_n^j}{y_n^j} + \frac{1}{y_n^j}\left(b^j(X_n) - \mathbb{E}[b^j(X_n)|\mathcal{F}_{n-\tau(\alpha_n)}]\right) \\ G_n^j &= \left(\mathbb{E}[A(X_n)|\mathcal{F}_{n-\tau(\alpha_n)}] - A\right)\frac{\theta_n^j}{y_n^j} + \frac{1}{y_n^j}\left(\mathbb{E}[b^j(X_n)|\mathcal{F}_{n-\tau(\alpha_n)}] - b^j\right).\end{aligned}$$

From equation 85

$$\theta_{n+1}^i = \sum_{j=1}^N \tilde{w}_n^{ij}\left[\theta_n^j + \alpha_n h^j(\theta_n, y_n) + \alpha_n M_n^j + \alpha_n G_n^j\right].$$

Let $h = [h^1, \cdots, h^N]^\top$, $M = [M^1, \cdots, M^N]^\top$ and $G = [G^1, \cdots, G^N]^\top$. Since

$$\begin{aligned}\mathbb{E}[M_n^j|\mathcal{F}_n] &= \left(\mathbb{E}[A(X_t)|\mathcal{F}_n] - \mathbb{E}[\mathbb{E}[A(X_t)|\mathcal{F}_{n-\tau(\alpha_n)}]|\mathcal{F}_n]\right)\frac{\theta_t^j}{y_t^j} \\ &\quad + \frac{1}{y_t^j}\left(\mathbb{E}[b^j(X_t)|\mathcal{F}_n] - \mathbb{E}[\mathbb{E}[b^j(X_t)|\mathcal{F}_{n-\tau(\alpha_n)}]|\mathcal{F}_n]\right) = 0\end{aligned}$$

and for all $n \geq \tau(\alpha_n)$

$$\begin{aligned}&\mathbb{E}[\|M_n\|_F^2|\mathcal{F}_n] = \sum_{j=1}^N \mathbb{E}[\|M_n^j\|_2^2|\mathcal{F}_n] \\ &= \sum_{j=1}^N \mathbb{E}[\|\left(A(X_n) - \mathbb{E}[A(X_n)|\mathcal{F}_{n-\tau(\alpha_n)}]\right)\frac{\theta_t^j}{y_t^j} + \frac{1}{y_t^j}\left(b^j(X_t) - \mathbb{E}[b^j(X_t)|\mathcal{F}_{n-\tau(\alpha_n)}]\right)\|_2^2|\mathcal{F}_n] \\ &\leq \sum_{j=1}^N \left(\frac{2A_{\max}+\alpha_0}{\beta}\|\theta_t^j\|_2 + \frac{2b_{\max}+\alpha_0}{\beta}\right)^2 \leq \frac{2(2A_{\max}+\alpha_0)^2}{\beta^2}\|\Theta_t\|_F^2 + \frac{2N}{\beta^2}(2b_{\max}+\alpha_0)^2,\end{aligned}$$

then $\{M_n\}$ is a martingale difference sequence satisfying $\mathbb{E}[\|M_n\|_F^2|\mathcal{F}_n] \leq \hat{C}(1 + \|\Theta_t\|_F)$, where $\hat{C} = \max\{\frac{2(2A_{\max}+\alpha_0)^2}{\beta^2}, \frac{2N}{\beta^2}(2b_{\max} + \alpha_0)^2\}$.

Define $h_c : \mathbb{R}^{N\times K} \times \mathbb{R}^N \to \mathbb{R}^{N\times K}$ as $h_c(x, y) = h(cx, y)c^{-1}$ with some $c \geq 1$. In addition, by using Lemma 15, define $\tilde{h}_c(z) : \mathbb{R}^K \to \mathbb{R}^K$ as $\tilde{h}_c(z) = h_c(\mathbf{1}_N \cdot z^\top, y_n)^\top \tilde{\pi}_n$, i.e.,

$$h_c(\Theta_n, y_n) = \begin{bmatrix} (A\frac{\theta_n^1}{y_n^1} + \frac{b^1}{y_n^1 c})^\top \\ \cdots \\ (A\frac{\theta_n^N}{y_n^N} + \frac{b^N}{y_n^N c})^\top \end{bmatrix}, \quad \tilde{h}_c(z) = Az + \sum_{i=1}^N \frac{b^i}{Nc}.$$

Then $\tilde{h}_c(z) \to \tilde{h}_\infty(z) = Az$ as $c \to \infty$ uniformly on compact sets. Let $\phi_c(z, t)$ and $\phi_\infty(z, t)$ denote the solutions of the ODE:

$$\dot{z}(t) = \tilde{h}_c(z(t)), \quad z(0) = z \tag{88}$$
$$\dot{z}(t) = \tilde{h}_\infty(z(t)) = Az(t), \quad z(0) = z$$

respectively. Furthermore, since the origin is the unique globally asymptotically stable equilibrium of the ODE, then we have the following lemma.

Lemma 16 *There exist constant $c_0 > 0$ and $T > 0$ such that for all initial conditions z with the sphere $\{z|\|z\|_2 \leq \frac{1}{N^{1/2}}\}$ and all $c \geq c_0$, we have $\|\phi_c(z, t)\|_2 < \frac{1-\kappa}{N^{1/2}}$ for $t \in [T, T + 1]$ for some $0 < \kappa < 1$.*

Proof of Lemma 16: Similar to the proof of Lemma 5 in Mathkar & Borkar (2016). ■

Define $t_0 = 0$, $t_n = \sum_{i=0}^n \alpha_n$, $n \geq 0$. Define $\bar{\Theta}(t), t \geq 0$ as $\bar{\Theta}(t_n) = \Theta_n$ with linear interpolation on each interval $[t_n, t_{n+1}]$. In addition, let $T_0 = 0$ and $T_{n+1} = \min\{t_m : t_m \geq T_n + T\}$ for all $n \geq 0$. Then, $T_{n+1} \in [T_n + T, T_n + T + \sup_n \alpha_n]$. Let $m(n)$ be the value such that $T_n = t_{m(n)}$ for any $n \geq 0$. Define the piecewise continuous trajectory $\hat{\Theta}(t) = \bar{\Theta}(t) \cdot r_n^{-1}$ for $t \in [T_n, T_{n+1})$, where $r_n = \max\{\|\bar{\Theta}(T_n)\|_F, 1\}$.

Lemma 17 *There exists a positive constant $C_{\hat{\theta}} < \infty$ such that $\sup_t \|\hat{\Theta}(t)\|_F < C_{\hat{\theta}}$.*

Proof of Lemma 17: First, we write the update of $\hat{\Theta}(t_k)$ for $k \in [m(n), m(n + 1))$

$$\hat{\Theta}(t_{k+1}) = \tilde{W}_{t_k}\left[\hat{\Theta}(t_k) + \alpha_{t_k}\begin{bmatrix} (\hat{\theta}^1(t_k))^\top/y_{t_k}^1 \\ \cdots \\ (\hat{\theta}^N(t_k))^\top/y_{t_k}^N \end{bmatrix} A(X_{t_k})^\top + \alpha_{t_k}\begin{bmatrix} (b^1(X_{t_k}))^\top/(y_{t_k}^1 r_n) \\ \cdots \\ (b^N(X_{t_k}))^\top/(y_{t_k}^N r_n) \end{bmatrix}\right]. \tag{89}$$

Since W_{t_k} is a column matrix, thus we have

$$\begin{aligned}
&\|\hat{\Theta}(t_{k+1})\|_\infty \\
&\leq \|\tilde{W}_{t_k}\|_\infty\left(\|\hat{\Theta}(t_k)\|_\infty + \alpha_{t_k}\left\|\begin{bmatrix} A(X_{t_k})\hat{\theta}^1(t_k)/y_{t_k}^1 \\ \cdots \\ A(X_{t_k})\hat{\theta}^N(t_k)/y_{t_k}^N \end{bmatrix}\right\|_\infty + \alpha_{t_k}\left\|\begin{bmatrix} \frac{b^1(X_{t_k})}{y_{t_k}^1 r_n} \\ \cdots \\ \frac{b^N(X_{t_k})}{y_{t_k}^N r_n} \end{bmatrix}\right\|_\infty\right) \\
&\leq \|\hat{\Theta}(t_k)\|_\infty + \frac{\alpha_{t_k}\sqrt{K}A_{\max}}{\beta}\|\hat{\Theta}(t_k)\|_\infty + \frac{\alpha_{t_k}\sqrt{K}b_{\max}}{\beta r_n} \\
&\leq \|\hat{\Theta}(t_{m(n)})\|_\infty + \sqrt{K}\sum_{l=0}^{k-m(n)} \frac{\alpha_{t_{k+l}}A_{\max}}{\beta}\|\hat{\Theta}(t_{k+l})\|_\infty + \frac{\alpha_{t_{k+l}}b_{\max}}{\beta r_n} \\
&\leq \sqrt{K} + \frac{(T + \sup_l \alpha_l)\sqrt{K}b_{\max}}{\beta} + \sum_{l=0}^{k-m(n)} \frac{\alpha_{t_{k+l}}\sqrt{K}A_{\max}}{\beta}\|\hat{\Theta}(t_{k+l})\|_\infty,
\end{aligned}$$

where we use the fact that $\|\hat{\Theta}(t_{m(n)})\|_F = 1$ and $r_n \geq 1$ in the last inequality. Therefore, by using discrete-time Grönwall inequality, we have

$$\sup_{m(n)\leq k<m(n+1)} \|\hat{\Theta}(t_{k+1})\|_\infty \leq \sqrt{K}(1+(T+\sup_l \alpha_l)b_{\max}) \exp\left\{\frac{A_{\max}\sqrt{K}}{\beta}(T+\sup_l \alpha_l)\right\}.$$

Since $T+\sup_l \alpha_l < \infty$, we have $\sup_{m(n)\leq k<m(n+1)} \|\hat{\Theta}(t_{k+1})\|_\infty < \infty$ for all n. By equivalence of vector norms, we further obtain that $\sup_t \|\hat{\Theta}(t)\|_F < \infty$. ■

For $n \geq 0$, let $z^n(t)$ denote the trajectory of $\dot{z} = \tilde{h}_c(z)$ with $c = r_n$ and $z^n(T_n) = \sum_{i=1}^N \tilde{\pi}^i_{T_n}\hat{\theta}_{T_n}$, for $[T_n, T_{n+1})$.

Lemma 18 $\lim_n \sup_{t\in[T_n,T_{n+1})} \|\hat{\Theta}_t - \mathbf{1}\otimes z^n(t)\| = 0.$

Proof of Lemma 18: From equation 85 and equation 89, for any $k \in [m(n), m(n+1))$, by Lemma 15, we have

$$\begin{aligned}
&\sum_{i=1}^N \tilde{\pi}^i_{n+1}\theta^i_{n+1} = \Theta^\top_{n+1}\tilde{\pi}_{n+1}\\
&= \left(\Theta_n + \alpha_n \begin{bmatrix} (A(X_n)\theta^1_n)^\top / y^1_n \\ \cdots \\ (A(X_n)\theta^N_n)^\top / y^N_n \end{bmatrix} + \alpha_n \begin{bmatrix} (b^1(X_n))^\top / y^1_n \\ \cdots \\ (b^N(X_n))^\top / y^N_n \end{bmatrix}\right)^\top \tilde{\pi}_n\\
&= \sum_{i=1}^N \tilde{\pi}^i_n\theta^i_n + \alpha_n \sum_{i=1}^N \tilde{\pi}^i_n (A(X_n)\theta^i_n / y^i_n + b^i(X_n)/y^i_n)\\
&= \sum_{i=1}^N \tilde{\pi}^i_n\theta^i_n + \frac{\alpha_n}{N} A(X_n)\sum_{i=1}^N \theta^i_n + \frac{\alpha_n}{N}\sum_{i=1}^N b^i(X_n).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\sum_{i=1}^N \tilde{\pi}^i_{t_{k+1}}\hat{\theta}^i_{t_{k+1}}\\
&= \sum_{i=1}^N \tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k} + \alpha_t \sum_{i=1}^N \tilde{\pi}^i_{t_k}(A(X_{t_k})\hat{\theta}^i_{t_k}/y^i_{t_k} + b^i(X_{t_k})/(y^i_{t_k} r_n))\\
&= \sum_{i=1}^N \tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k} + \alpha_{t_k}\left(A(X_{t_k})\sum_{i=1}^N \tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k} + \frac{1}{Nr_n}\sum_{i=1}^N b^i(X_{t_k})\right)\\
&\quad + \alpha_{t_k}\frac{A(X_{t_k})}{N}\sum_{i=1}^N\left(\hat{\theta}^i_{t_k} - \sum_{i=1}^N \tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k}\right)\\
&= \sum_{i=1}^N \tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k} + \alpha_{t_k}\left(A\sum_{i=1}^N \tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k} + \frac{1}{Nr_n}\sum_{i=1}^N b^i\right) + \alpha_{t_k}\frac{A(X_{t_k})}{N}\sum_{i=1}^N\left(\hat{\theta}^i_{t_k} - \sum_{i=1}^N \tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k}\right)\\
&\quad + \alpha_{t_k}\left(A(X_{t_k}) - \mathbb{E}[A(X_{t_k})|\mathcal{F}_{t_k-\tau(\alpha_{t_k})}]\right)\sum_{i=1}^N \tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k}\\
&\quad + \frac{\alpha_{t_k}}{Nr_n}\sum_{i=1}^N\left(b^i(X_{t_k}) - \mathbb{E}[b^i(X_{t_k})|\mathcal{F}_{t_k-\tau(\alpha_{t_k})}]\right)\\
&\quad + \alpha_{t_k}\left(\left(\mathbb{E}[A(X_{t_k})|\mathcal{F}_{t_k-\tau(\alpha_{t_k})}] - A\right)\sum_{i=1}^N \tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k} + \frac{1}{Nr_n}\sum_{i=1}^N\left(\mathbb{E}[b^i(X_{t_k})|\mathcal{F}_{t_k-\tau(\alpha_{t_k})}] - b^i\right)\right).
\end{aligned}$$

Let

$$\hat{M}_{t_k} = \Big(A(X_t) - \mathbb{E}[A(X_{t_k})|\mathcal{F}_{t_k-\tau(\alpha_{t_k})}]\Big)\sum_{i=1}^{N}\tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k} + \frac{1}{Nr_n}\sum_{i=1}^{N}\Big(b^i(X_{t_k}) - \mathbb{E}[b^i(X_{t_k})|\mathcal{F}_{t_k-\tau(\alpha_{t_k})}]\Big)$$

$$\hat{G}_{t_k} = \Big(\mathbb{E}[A(X_{t_k})|\mathcal{F}_{t_k-\tau(\alpha_{t_k})}] - A\Big)\sum_{i=1}^{N}\tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k} + \frac{1}{Nr_n}\sum_{i=1}^{N}\Big(\mathbb{E}[b^i(X_{t_k})|\mathcal{F}_{t_k-\tau(\alpha_{t_k})}] - b^i\Big) + \frac{A(X_{t_k})}{N}\sum_{i=1}^{N}\left(\hat{\theta}^i_{t_k} - \sum_{i=1}^{N}\tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k}\right).$$

It is easy to verify that $\{\hat{M}_{t_k}\}$ is a martingale difference sequence satisfying $\mathbb{E}[\|\hat{M}_{t_k}\|_2^2|\mathcal{F}_{t_k}] \le \bar{C}(1 + \|\sum_{i=1}^{N}\tilde{\pi}^i_{t_k}\hat{\theta}^i_{t_k}\|_2^2)$ for some $\bar{C} \le \infty$. In addition, we have

$$\hat{\theta}^i_{t_k} - \sum_{j=1}^{N}\tilde{\pi}^j_{t_k}\hat{\theta}^j_{t_k} = \sum_{j=1}^{N}(\tilde{w}^{ij}_{t_s:t_k} - \tilde{\pi}^j_{t_s})\hat{\theta}^j_{t_s} + \sum_{r=s+1}^{k}\alpha_{t_r}\sum_{i=1}^{N}(\tilde{w}^{ij}_{t_r:t_k} - \tilde{\pi}^j_{t_r})(A(X_{t_r})\hat{\theta}^j_{t_r}/y^j_{t_r} + b^j(X_{t_r}/y^j_{t_r}).$$

Since $\{\mathbb{G}_t\}$ is uniformly strongly connected, then for any $s \ge 0$, $W_{s:t}$ converges to $\mathbf{1}\pi_s^\top$ exponentially fast as $t \to \infty$ and there exist a finite positive constant C and a constant $0 \le \lambda < 1$ such that

$$|\tilde{w}^{ij}_{s:t} - \tilde{\pi}^j_s| \le C\lambda^{t-s}$$

for all $i, j \in \mathcal{V}$ and $s \ge 0$. Then,for any $k \in [m(n), m(n+1))$, we have

$$\begin{aligned}
&\|\hat{\theta}^i_{t_k} - \sum_{j=1}^{N}\tilde{\pi}^j_{t_k}\hat{\theta}^j_{t_k}\|_2 \\
&\le \sum_{j=1}^{N}\|\tilde{w}^{ij}_{t_{m(n)}:t_k} - \tilde{\pi}^j_{t_{m(n)}}\|_2\|\hat{\theta}^j_{t_{m(n)}}\|_2 + \sum_{r=m(n)+1}^{k}\alpha_{t_r}\sum_{i=1}^{N}\|\tilde{w}^{ij}_{t_r:t_k} - \tilde{\pi}^j_{t_r}\|_2\frac{A_{\max}\|\hat{\theta}^j_{t_r}\|_2 + b_{\max}}{\beta} \\
&\le \sum_{j=1}^{N}C\lambda^{t_k - t_{m(n)}}\|\hat{\theta}^j_{t_{m(n)}}\|_2 + \sum_{r=m(n)+1}^{k}\alpha_{t_r}\sum_{i=1}^{N}C\lambda^{t_k-t_r}(\frac{A_{\max}\|\hat{\theta}^j_{t_r}\|_2 + b_{\max}}{\beta}) \\
&\le NC\lambda^{t_k - t_{m(n)}} + \frac{\alpha_{t_{m(n)}}NC}{1-\lambda}\frac{A_{\max}C_{\hat{\theta}} + b_{\max}}{\beta},
\end{aligned}$$

where in the last inequality, we use the fact that for all $n \ge 0$, we have $\|\hat{\Theta}(t_{m(n)})\|_F = 1$, $\alpha_{n+1} \le \alpha_n$ and the boundedness of $\|\hat{\Theta}_n\|_F$ from Lemma 17. Since $\alpha_{t_k} \to 0$ as $k \to \infty$, then

$$\lim_{k\to\infty}\|\hat{\theta}^i_{t_k} - \sum_{j=1}^{N}\tilde{\pi}^j_{t_k}\hat{\theta}^j_{t_k}\|_2 = 0,$$

which implies that

$$\lim_{k\to\infty}\left\|\frac{A(X_{t_k})}{N}\sum_{i=1}^{N}(\hat{\theta}^i_{t_k} - \sum_{j=1}^{N}\tilde{\pi}^j_{t_k}\hat{\theta}^j_{t_k})\right\|_2 = 0.$$

Then,

$$\lim_{k\to\infty}\|\hat{G}_{t_k}\|_2 \le \lim_{k\to\infty}\alpha_{t_k}(\|\sum_{j=1}^{N}\tilde{\pi}^j_{t_k}\hat{\theta}^j_{t_k}\|_2 + 1) + \lim_{k\to\infty}\left\|\frac{A(X_{t_k})}{N}\sum_{i=1}^{N}(\hat{\theta}^i_{t_k} - \sum_{j=1}^{N}\tilde{\pi}^j_{t_k}\hat{\theta}^j_{t_k})\right\|_2 = 0.$$

Therefore, by Corollary 8 and Theorem 9 in Chapter 6 of Borkar (2008), we obtain that $\sum_{i=1}^{N}\tilde{\pi}_{t_k}^i\hat{\theta}_{t_k}^i \to z^n(t)$ as $n\to\infty$, namely $k\to\infty$. Furthermore, we obtain that $\hat{\theta}_{t_{k+1}}^i \to z^n(t)$ as $n\to\infty$ for all $i\in\mathcal{V}$, which concludes the proof following Theorem 2 in Chapter 2 of Borkar (2008). ■

Lemma 19 *The sequence $\{\Theta_n\}$ generated from equation 87 is bounded almost surely, i.e., $C_\theta = \sup_n \|\Theta_n\|_F < \infty$ almost surely.*

Proof of Lemma 19: In order to prove this lemma, we need to show that $\sup_n \|\bar{\Theta}(T_n)\|_F < \infty$ first. If this does not hold, there will exist a sequence $T_{n_1}, T_{n_2}, \cdots$ such that $\|\hat{\Theta}(T_{n_k})\|_F \to \infty$, i.e., $r_{n_k}\to\infty$. If $r_n > c_0$ and $\|\hat{\Theta}(T_n)\|_F = 1$, then $\|z^n(T_n)\|_2 = \|\sum_{i=1}^N \tilde{\pi}_{T_n}\hat{\theta}_{T_n}^i\|_2 \le N^{-1/2}$. Using Lemma 16, we have $\|\mathbf{1}_N\cdot(z^n(T_{n+1}^-))^\top\|_F = N^{1/2}\|z^n(T_{n+1}^-)\|_2 \le 1-\kappa$. In addition, using Lemma 18, there exists a constant $0<\kappa'<\kappa$ such that $\|\hat{\Theta}(T_{n+1}^-)\|_F < 1-\kappa'$. Hence for $r_n > c_0$ and n sufficiently large,

$$\frac{\|\bar{\Theta}(T_{n+1})\|_F}{\|\bar{\Theta}(T_n)\|_F} = \frac{\|\hat{\Theta}(T_{n+1}^-)\|_F}{\|\hat{\Theta}(T_n)\|_F} \le 1-\kappa'.$$

It shows that if $\|\bar{\Theta}(T_n)\|_F > c_0$, $\|\bar{\Theta}(T_k)\|_F$ for all $k\ge n$ falls back to the ball of radius c_0 at an exponential rate.

Thus, if $\|\bar{\Theta}(T_n)\|_F > c_0$, then $\|\bar{\Theta}(T_{n-1})\|_F$ is either greater than $\|\bar{\Theta}(T_n)\|_F$ or is inside the ball of radius c_0. Since we assume $r_{n_k}\to\infty$, then we can find a time T_n such that $\|\bar{\Theta}(T_n)\|_F < c_0$ and $\|\bar{\Theta}(T_{n+1})\|_F = \infty$. However, by using discrete-time Grönwall inequality, we have

$$\begin{aligned}
\|\bar{\Theta}(T_{n+1})\|_\infty &\le \|\bar{\Theta}(T_{n+1}-1)\|_\infty + \alpha_{T_{n+1}-1}\frac{\sqrt{K}A_{\max}}{\beta}\|\bar{\Theta}(T_{n+1}-1)\|_\infty + \alpha_{T_{n+1}-1}\sqrt{K}\frac{b_{\max}}{\beta}\\
&\le \|\bar{\Theta}(T_n)\|_\infty + \sqrt{K}\sum_{s=0}^{T_{n+1}-T_n}\alpha_{T_n+s}\frac{A_{\max}}{\beta}\|\bar{\Theta}(T_n+s)\|_\infty + \alpha_{T_n+s}\frac{b_{\max}}{\beta}\\
&\le \sqrt{K}c_0 + \sqrt{K}(T+\sup_n \alpha_n)\frac{b_{\max}}{\beta} + \frac{\sqrt{K}A_{\max}}{\beta}\sum_{s=0}^{T_{n+1}-T_n}\alpha_{T_n+s}\|\bar{\Theta}(T_n+s)\|_\infty\\
&\le \sqrt{K}(c_0 + (T+\sup_n\alpha_n)\frac{b_{\max}}{\beta})\exp\left\{(T+\sup_n\alpha_n)\frac{\sqrt{K}A_{\max}}{\beta}\right\},
\end{aligned}$$

which implies that $\|\bar{\Theta}(T_{n+1})\|_F$ can be bounded if $\|\bar{\Theta}(T_n)\|_F < c_0$. This leads to a contradiction. Moreover, let $C_{\bar{\theta}} = \sup_n \|\bar{\Theta}(T_n)\|_F < \infty$, then $C_\theta = \sup_n\|\Theta_n\|_F \le C_{\bar{\theta}}C_{\hat{\theta}} < \infty$. ■

Recall the update of $\tilde{\theta}_t^i$ in equation 9:

$$\tilde{\theta}_{t+1}^i = \sum_{j=1}^N \hat{w}_t^{ij}\left[\tilde{\theta}_t^j + \alpha_t\left(A(X_t)\theta_t^j + b^j(X_t)\right)\right].$$

From the definition that $\langle\tilde{\theta}\rangle_t = \frac{1}{N}\sum_{i=1}^N\tilde{\theta}_t^i$ and $\langle\theta\rangle_t = \frac{1}{N}\sum_{i=1}^N\theta_t^i$, we have

$$\begin{aligned}
\langle\tilde{\theta}\rangle_{t+1} &= \langle\tilde{\theta}\rangle_t + \alpha_t A(X_t)\langle\theta\rangle_t + \frac{\alpha_t}{N}\sum_{i=1}^N b^i(X_t)\\
&= \langle\tilde{\theta}\rangle_t + \alpha_t A(X_t)\langle\tilde{\theta}\rangle_t + \frac{\alpha_t}{N}\sum_{i=1}^N b^i(X_t) + \alpha_t\rho_t, \qquad (90)
\end{aligned}$$

where $\rho_t = A(X_t)\langle\theta\rangle_t - A(X_t)\langle\tilde{\theta}\rangle_t$. From Lemma 19, we have $\|\langle\theta\rangle_t\|_2 \le \max_{i\in\mathcal{V}}\|\theta_t^i\|_2 \le C_\theta$ for all $t\ge 0$, which implies that $\|\langle\tilde{\theta}\rangle_t\|_2 \le NC_\theta$ and

$$\mu_t = \|\rho_t\|_2 = \left\|A(X_t)\langle\theta\rangle_t - A(X_t)\langle\tilde{\theta}\rangle_t\right\|_2 \le \mu_{\max},$$

where $\mu_{\max} = (N+1)A_{\max}C_\theta$.

Lemma 20 *Suppose that Assumptions 2 and 5 hold and $\{\mathbb{G}_t\}$ is uniformly strongly connected by sub-sequences of length L. Let $\epsilon_1 = \inf_{t\geq 0}\min_{i\in\mathcal{V}}(\hat{W}_t\cdots\hat{W}_0\mathbf{1}_N)^i$. For all $t\geq 0$ and $i\in\mathcal{V}$,*

$$\|\theta_{t+1}^i - \langle\tilde{\theta}\rangle_t\|_2 \leq \frac{8}{\epsilon_1}\bar{\epsilon}^t\|\sum_{i=1}^N \tilde{\theta}_0^i + \alpha_0 A(X_0)\theta_0^i + \alpha_0 b^i(X_0)\|_2 + \frac{8}{\epsilon_1}\frac{A_{\max}C_\theta + b_{\max}}{1-\bar{\epsilon}}\left(\alpha_0\bar{\epsilon}^{t/2} + \alpha_{\lceil\frac{t}{2}\rceil}\right) + \alpha_t A_{\max}C_\theta + \alpha_t b_{\max},$$

where $\epsilon_1 > 0$ and $\bar{\epsilon}\in(0,1)$ satisfy $\epsilon_1 \geq \frac{1}{N^{NL}}$ and $\bar{\epsilon}\leq(1-\frac{1}{N^{NL}})^{1/L}$.

Proof of Lemma 20: Since $\epsilon_1 = \inf_{t\geq 0}\min_{i\in\mathcal{V}}(\hat{W}_t\cdots\hat{W}_0\mathbf{1}_N)^i$ and all weight matrices $\hat{W}_s$ are column stochastic matrices for all $s\geq 0$, from Corollary 2 (b) in Nedić & Olshevsky (2015), we know that $\epsilon_1 \leq \frac{1}{N^{NL}}$. If the weight matrices are doubly stochastic matrices, then $\epsilon_1 = 1$.

From Assumption 2 and Lemma 19, we know that $\|A(X_t)\theta_t^i + b^i(X_t)\|_2 \leq A_{\max}C_\theta + b_{\max}$. Then, by using Lemma 1 in Nedić & Olshevsky (2015), for all $t\geq 0$ and $i\in\mathcal{V}$ we have

$$\begin{aligned}
&\|\theta_{t+1}^i - \langle\tilde{\theta}\rangle_t - \alpha_t A(X_t)\langle\theta\rangle_t - \frac{\alpha_t}{N}\sum_{i=1}^N b^i(X_t)\|_2 \\
&\leq \frac{8}{\epsilon_1}(\bar{\epsilon}^t\|\sum_{i=1}^N \tilde{\theta}_0^i + \alpha_0 A(X_0)\theta_0^i + \alpha_0 b^i(X_0)\|_2 + \sum_{s=0}^t \bar{\epsilon}^{t-s}\alpha_s(A_{\max}C_\theta + b_{\max})) \\
&\leq \frac{8}{\epsilon_1}\bar{\epsilon}^t\|\sum_{i=1}^N \tilde{\theta}_0^i + \alpha_0 A(X_0)\theta_0^i + \alpha_0 b^i(X_0)\|_2 \\
&\quad + \frac{8}{\epsilon_1}(A_{\max}C_\theta + b_{\max})\left(\sum_{s=0}^{\lfloor\frac{t}{2}\rfloor}\bar{\epsilon}^{t-s}\alpha_s + \sum_{s=\lceil\frac{t}{2}\rceil}^t \bar{\epsilon}^{t-s}\alpha_s\right) \\
&\leq \frac{8}{\epsilon_1}\bar{\epsilon}^t\|\sum_{i=1}^N \tilde{\theta}_0^i + \alpha_0 A(X_0)\theta_0^i + \alpha_0 b^i(X_0)\|_2 + \frac{8}{\epsilon_1}\frac{A_{\max}C_\theta + b_{\max}}{1-\bar{\epsilon}}\left(\alpha_0\bar{\epsilon}^{t/2} + \alpha_{\lceil\frac{t}{2}\rceil}\right),
\end{aligned}$$

which implies that

$$\begin{aligned}
&\|\theta_{t+1}^i - \langle\tilde{\theta}\rangle_t\|_2 \\
&\leq \|\theta_{t+1}^i - \langle\tilde{\theta}\rangle_t - \alpha_t A(X_t)\langle\theta\rangle_t - \frac{\alpha_t}{N}\sum_{i=1}^N b^i(X_t)\|_2 + \alpha_t\|A(X_t)\langle\theta\rangle_t + \frac{1}{N}\sum_{i=1}^N b^i(X_t)\|_2 \\
&\leq \frac{8}{\epsilon_1}\bar{\epsilon}^t\|\sum_{i=1}^N \tilde{\theta}_0^i + \alpha_0 A(X_0)\theta_0^i + \alpha_0 b^i(X_0)\|_2 + \frac{8}{\epsilon_1}\frac{A_{\max}C_\theta + b_{\max}}{1-\bar{\epsilon}}\left(\alpha_0\bar{\epsilon}^{t/2} + \alpha_{\lceil\frac{t}{2}\rceil}\right) \\
&\quad + \alpha_t A_{\max}C_\theta + \alpha_t b_{\max}.
\end{aligned}$$

This completes the proof. ■

Lemma 21 $\lim_{t\to\infty}\mu_t = \lim_{t\to\infty}\|\rho_t\|_2 = 0$ *and* $\lim_{t\to\infty}\frac{\sum_{k=0}^t \mu_k}{t+1} = \lim_{t\to\infty}\frac{\sum_{k=0}^t\|\rho_k\|_2}{t+1} = 0$.

Proof of Lemma 21: From Lemma 20, we have

$$\begin{aligned}
\mu_t = \|\rho_t\|_2 &= \left\|A(X_t)\langle\theta\rangle_t - A(X_t)\langle\tilde{\theta}\rangle_t\right\|_2 \\
&\leq \frac{8A_{\max}}{\epsilon_1}\bar{\epsilon}^t\|\tilde{\Theta}_0\|_1 + \frac{8A_{\max}}{\epsilon_1}\frac{N\sqrt{K}(A_{\max}C_\theta + b_{\max})}{1-\bar{\epsilon}}\left(\alpha_0\bar{\epsilon}^{t/2} + \alpha_{\lceil\frac{t}{2}\rceil}\right).
\end{aligned}$$

Since $\bar{\epsilon}\in(0,1)$, then $\lim_{t\to\infty}\|\rho_t\|_2 = 0$. Next, we will prove that $\lim_{t\to\infty}\frac{1}{t+1}\sum_{k=0}^t\|\rho_k\|_2 = 0$. For any positive constant $c > 0$, there exists a positive integer $T(c)$, depending on c, such that

$\forall t \geq T(c)$, we have $\|\rho_t\|_2 < c$. Thus,

$$\frac{1}{t}\sum_{k=0}^{t-1}\|\rho_k\|_2 = \frac{1}{t}\sum_{k=0}^{T(c)}\|\rho_k\|_2 + \frac{1}{t}\sum_{k=T(c)+1}^{t-1}\|\rho_k\|_2 \leq \frac{1}{t}\sum_{k=0}^{T(c)}\|\rho_k\|_2 + \frac{t-1-T(c)}{t}c.$$

Let $t \to \infty$ on both sides of the above inequality. Then, we have

$$\lim_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t-1}\|\rho_k\|_2 \leq \lim_{t\to\infty}\frac{1}{t}\sum_{k=0}^{T(c)}\|\rho_k\|_2 + \lim_{t\to\infty}\frac{t-1-T(c)}{t}c = c.$$

Since the above argument holds for arbitrary positive c, then $\lim_{t\to\infty}\frac{1}{t+1}\sum_{k=0}^{t}\|\rho_k\|_2 = 0$. ■

Lemma 22 *Suppose that Assumptions 2 and 3 hold. When the step-size α_t and corresponding mixing time $\tau(\alpha_t)$ satisfy $0 < \alpha_t\tau(\alpha_t) < \frac{\log 2}{A_{\max}}$, we have for any $t \geq \bar{T}$,*

$$\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \leq 2A_{\max}\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k + 2(b_{\max}+\mu_{\max})\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k, \tag{91}$$

$$\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \leq 6A_{\max}\|\langle\tilde{\theta}\rangle_t\|_2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k + 5(b_{\max}+\mu_{\max})\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k, \tag{92}$$

$$\begin{aligned}\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 &\leq 72\alpha_{t-\tau(\alpha_t)}^2\tau^2(\alpha_t)A_{\max}^2\|\langle\tilde{\theta}\rangle_t\|_2^2 + 50\alpha_{t-\tau(\alpha_t)}^2\tau^2(\alpha_t)(b_{\max}+\mu_{\max})^2\\ &\leq 8\|\langle\tilde{\theta}\rangle_t\|_2^2 + \frac{6(b_{\max}+\mu_{\max})^2}{A_{\max}^2}.\end{aligned} \tag{93}$$

Proof of Lemma 22: From the update of $\langle\tilde{\theta}\rangle_t$ in equation 90:

$$\langle\tilde{\theta}\rangle_{t+1} = \langle\tilde{\theta}\rangle_t + \alpha_t A(X_t)\langle\tilde{\theta}\rangle_t + \frac{\alpha_t}{N}\sum_{i=1}^{N}b^i(X_t) + \alpha_t\rho_t.$$

Then, we have

$$\|\langle\tilde{\theta}\rangle_{t+1}\|_2 \leq (1+\alpha_t A_{\max})\|\langle\tilde{\theta}\rangle_t\|_2 + \alpha_t b_{\max} + \alpha_t\mu_{\max}.$$

For all $u \in [t-\tau(\alpha_t), t]$, we have

$$\begin{aligned}\|\langle\tilde{\theta}\rangle_u\|_2 &\leq \Pi_{k=t-\tau(\alpha_t)}^{u-1}(1+\alpha_k A_{\max})\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2\\ &\quad + (b_{\max}+\mu_{\max})\sum_{k=t-\tau(\alpha_t)}^{u-1}\alpha_k\Pi_{l=k+1}^{u-1}(1+\alpha_l A_{\max})\\ &\leq \exp\{\sum_{k=t-\tau(\alpha_t)}^{u-1}\alpha_k A_{\max}\}\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2\\ &\quad + (b_{\max}+\mu_{\max})\sum_{k=t-\tau(\alpha_t)}^{u-1}\alpha_k\exp\{\sum_{l=k+1}^{u-1}\alpha_l A_{\max}\}\\ &\leq \exp\{\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)A_{\max}\}\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2\\ &\quad + (b_{\max}+\mu_{\max})\sum_{k=t-\tau(\alpha_t)}^{u-1}\alpha_k\exp\{\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)A_{\max}\}\\ &\leq 2\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 + 2(b_{\max}+\mu_{\max})\sum_{k=t-\tau(\alpha_t)}^{u-1}\alpha_k,\end{aligned}$$

where we use $\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)A_{\max} \leq \log 2 < \frac{1}{3}$ in the last inequality. Thus, for all $t \geq \bar{T}$, we can get equation 91 as follows:

$$\begin{aligned}
&\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \\
\leq& \sum_{k=t-\tau(\alpha_t)}^{t-1} \|\langle\tilde{\theta}\rangle_{k+1} - \langle\tilde{\theta}\rangle_k\|_2 \\
\leq& A_{\max} \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \|\langle\tilde{\theta}\rangle_k\|_2 + (b_{\max} + \mu_{\max}) \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
\leq& A_{\max} \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \left(2\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 + 2(b_{\max} + \mu_{\max}) \sum_{l=t-\tau(\alpha_t)}^{k-1} \alpha_l \right) \\
&+ (b_{\max} + \mu_{\max}) \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
\leq& 2A_{\max}\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k + \left(2A_{\max}\tau(\alpha_t)\alpha_{t-\tau(\alpha_t)} + 1\right)(b_{\max} + \mu_{\max}) \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
\leq& 2A_{\max}\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k + \frac{5}{3}(b_{\max} + \mu_{\max}) \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
\leq& 2A_{\max}\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k + 2(b_{\max} + \mu_{\max}) \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k.
\end{aligned}$$

Moreover, by using the above inequality, we can get equation 92 for all $t \geq \bar{T}$ as follows:

$$\begin{aligned}
&\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \\
\leq& 2A_{\max}\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k + \frac{5}{3}(b_{\max} + \mu_{\max}) \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
\leq& 2A_{\max}\tau(\alpha_t)\alpha_{t-\tau(\alpha_t)}\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 + 2A_{\max}\|\langle\tilde{\theta}\rangle_t\|_2 \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&+ \frac{5}{3}(b_{\max} + \mu_{\max}) \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
\leq& 6A_{\max}\|\langle\tilde{\theta}\rangle_t\|_2 \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k + 5(b_{\max} + \mu_{\max}) \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k.
\end{aligned}$$

Next, by using equation 92 and the inequality $(x+y)^2 \leq 2x^2+y^2$ for all x, y, we can get equation 93 as follows:

$$\begin{aligned}
\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 &\leq 72A_{\max}^2\|\langle\tilde{\theta}\rangle_t\|_2^2 (\sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k)^2 + 50(b_{\max} + \mu_{\max})^2 (\sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k)^2 \\
&\leq 72\alpha_{t-\tau(\alpha_t)}^2 \tau^2(\alpha_t) A_{\max}^2 \|\langle\tilde{\theta}\rangle_t\|_2^2 + 50\alpha_{t-\tau(\alpha_t)}^2 \tau^2(\alpha_t)(b_{\max} + \mu_{\max})^2 \\
&\leq 8\|\langle\tilde{\theta}\rangle_t\|_2^2 + \frac{6(b_{\max} + \mu_{\max})^2}{A_{\max}^2},
\end{aligned}$$

where we use $\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)A_{\max} < \frac{1}{3}$ in the last inequality. ■

Lemma 23 *Suppose that Assumptions 2–5 hold and $\{\mathbb{G}_t\}$ is uniformly strongly connected by sub-sequences of length L. When $0 < \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t) < \frac{\log 2}{A_{\max}}$, we have for any $t \geq \bar{T}$,*

$$\begin{aligned}
&|\mathbf{E}[(\langle\tilde{\theta}\rangle_t - \theta^*)^\top (P + P^\top)(A(X_t)\langle\tilde{\theta}\rangle_t + B(X_t)^\top \pi_{t+1} - A\langle\tilde{\theta}\rangle_t - b) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\left(72 + 456A_{\max}^2 + 84A_{\max}b_{\max} + 72A_{\max}\mu_{\max}\right)\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\quad + \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\Bigg[2 + 4\|\theta^*\|_2^2 + 48\frac{(b_{\max}+\mu_{\max})^2}{A_{\max}^2} + 152\left(b_{\max}+\mu_{\max}+A_{\max}\|\theta^*\|_2\right)^2 \\
&\quad + 12A_{\max}b_{\max} + 48A_{\max}(b_{\max}+\mu_{\max})(\frac{b_{\max}+\mu_{\max}}{A_{\max}} + 1)^2 + 87(b_{\max}+\mu_{\max})^2\Bigg].
\end{aligned}$$

Proof of Lemma 23: Note that for all $t \geq \bar{T}$, we have

$$\begin{aligned}
&\mathbf{E}[(\langle\tilde{\theta}\rangle_t - \theta^*)^\top (P + P^\top)(A(X_t)\langle\tilde{\theta}\rangle_t + \frac{1}{N}B(X_t)^\top \mathbf{1}_N - A\langle\tilde{\theta}\rangle_t - b) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq |\mathbf{E}[(\langle\tilde{\theta}\rangle_t - \theta^*)^\top (P + P^\top)(A(X_t) - A)\langle\tilde{\theta}\rangle_t \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\quad + |\mathbf{E}[(\langle\tilde{\theta}\rangle_t - \theta^*)^\top (P + P^\top)(\frac{1}{N}B(X_t)^\top \mathbf{1}_N - b) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq |\mathbf{E}[(\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top (P + P^\top)(A(X_t) - A)\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} \mid \mathcal{F}_{t-\tau(\alpha_t)}]| && (94) \\
&\quad + |\mathbf{E}[(\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top (P + P^\top)(A(X_t) - A)(\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| && (95) \\
&\quad + |\mathbf{E}[(\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)})^\top (P + P^\top)(A(X_t) - A)\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} \mid \mathcal{F}_{t-\tau(\alpha_t)}]| && (96) \\
&\quad + |\mathbf{E}[(\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)})^\top (P + P^\top)(A(X_t) - A)(\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| && (97) \\
&\quad + |\mathbf{E}[(\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)})^\top (P + P^\top)(\frac{1}{N}B(X_t)^\top \mathbf{1}_N - b) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| && (98) \\
&\quad + |\mathbf{E}[(\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top (P + P^\top)(\frac{1}{N}B(X_t)^\top \mathbf{1}_N - b) \mid \mathcal{F}_{t-\tau(\alpha_t)}]|. && (99)
\end{aligned}$$

By using the mixing time in Assumption 3, we can get the bound for equation 94 and equation 99 for all $t \geq \bar{T}$:

$$\begin{aligned}
&|\mathbf{E}[(\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top (P + P^\top)(A(X_t) - A)\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq |(\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top (P + P^\top)\mathbf{E}[A(X_t) - A \mid \mathcal{F}_{t-\tau(\alpha_t)}]\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}| \\
&\leq 2\alpha_t\gamma_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \theta^*\|_2\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq \alpha_t\gamma_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \theta^*\|_2^2 + \|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq \alpha_t\gamma_{\max}\mathbf{E}[2\|\theta^*\|_2^2 + 3\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq 6\alpha_t\gamma_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] + 6\alpha_t\gamma_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] + 2\alpha_t\gamma_{\max}\|\theta^*\|_2^2 \\
&\leq 54\alpha_t\gamma_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] + 36\alpha_t\gamma_{\max}\frac{(b_{\max}+\mu_{\max})^2}{A_{\max}^2} + 2\alpha_t\gamma_{\max}\|\theta^*\|_2^2, && (100)
\end{aligned}$$

where in the last inequality, we use equation 91 from Lemma 22.

$$\begin{aligned}
&|\mathbf{E}[(\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top (P + P^\top)(\frac{1}{N}B(X_t)^\top \mathbf{1}_N - b) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq |(\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top (P + P^\top)\frac{1}{N}\sum_{i=1}^N \mathbf{E}[b^i(X_t) - b^i \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq 2\gamma_{\max}\alpha_t\mathbf{E}[\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \theta^*\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq 2\gamma_{\max}\alpha_t\left(1 + \frac{1}{2}\mathbf{E}[\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] + \frac{1}{2}\|\theta^*\|_2^2\right) \\
&\leq 2\gamma_{\max}\alpha_t\left(1 + \mathbf{E}[\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \langle\tilde{\theta}\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] + \mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] + \|\theta^*\|_2^2\right) \\
&\leq 2\gamma_{\max}\alpha_t\left(1 + 9\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] + 6\frac{(b_{\max}+\mu_{\max})^2}{A_{\max}^2} + \|\theta^*\|_2^2\right), && (101)
\end{aligned}$$

where in the last inequality we use equation 91.

Next, by using Assumption 2, equation 91 and equation 93, we have

$$
\begin{aligned}
&|\mathbf{E}[(\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \theta^*)^\top (P + P^\top)(A(X_t) - A)(\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq 4\gamma_{\max} A_{\max} \mathbf{E}[\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} - \theta^*\|_2 \|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq 4\gamma_{\max} A_{\max} \mathbf{E}[\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\quad + 4\gamma_{\max} A_{\max} \|\theta^*\|_2 \mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq 8\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 8\gamma_{\max} A_{\max} (b_{\max} + \mu_{\max}) \|\theta^*\|_2 \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 8\gamma_{\max} A_{\max}^2 \left(\frac{b_{\max} + \mu_{\max}}{A_{\max}} + \|\theta^*\|_2 \right) \mathbf{E}[\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\leq 12\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 8\gamma_{\max} A_{\max}^2 \left(\frac{b_{\max} + \mu_{\max}}{A_{\max}} + \|\theta^*\|_2 \right)^2 \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\leq 24\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 24\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 8\gamma_{\max} A_{\max}^2 \left(\frac{b_{\max} + \mu_{\max}}{A_{\max}} + \|\theta^*\|_2 \right)^2 \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\leq 216\gamma_{\max} A_{\max}^2 \mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 152\gamma_{\max} \left(b_{\max} + \mu_{\max} + A_{\max} \|\theta^*\|_2 \right)^2 \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k.
\end{aligned} \tag{102}
$$

In additional, as for the bound of equation 96, by using equation 91 and equation 93, we have

$$\begin{aligned}
&|\mathbf{E}[(\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)})^\top (P+P^\top)(A(X_t)-A)\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)} \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq 4\gamma_{\max}A_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq 8\gamma_{\max}A_{\max}\mathbf{E}[A_{\max}\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 + (b_{\max}+\mu_{\max})\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\leq 4\gamma_{\max}A_{\max}(2A_{\max}+b_{\max}+\mu_{\max})\mathbf{E}[\|\langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 4\gamma_{\max}A_{\max}(b_{\max}+\mu_{\max}) \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\leq 8\gamma_{\max}A_{\max}(2A_{\max}+b_{\max}+\mu_{\max})\mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 8\gamma_{\max}A_{\max}(2A_{\max}+b_{\max}+\mu_{\max})\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 4\gamma_{\max}A_{\max}(b_{\max}+\mu_{\max}) \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\leq 72\,\gamma_{\max}A_{\max}(2A_{\max}+b_{\max}+\mu_{\max})\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + 48\gamma_{\max}A_{\max}(b_{\max}+\mu_{\max})(\frac{b_{\max}+\mu_{\max}}{A_{\max}}+1)^2 \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k. \qquad (103)
\end{aligned}$$

Moreover, by using equation 93, we can get the bound for equation 97 as follows:

$$\begin{aligned}
&|\mathbf{E}[(\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)})^\top (P+P^\top)(A(X_t)-A)(\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq 4\gamma_{\max}A_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq 4\gamma_{\max}A_{\max}\mathbf{E}[72A_{\max}^2\|\langle\tilde{\theta}\rangle_t\|_2^2 + 50(b_{\max}+\mu_{\max})^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \left(\sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k\right)^2 \\
&\leq 96A_{\max}^2\gamma_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k + 67(b_{\max}+\mu_{\max})^2\gamma_{\max} \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k. \qquad (104)
\end{aligned}$$

Finally, we can get the bound of equation 98 by using equation 92:

$$\begin{aligned}
&|\mathbf{E}[(\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)})(P+P^\top)(\frac{1}{N}B(X_t)^\top \mathbf{1}_N - b) \mid \mathcal{F}_{t-\tau(\alpha_t)}]| \\
&\leq 4\gamma_{\max}b_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \langle\tilde{\theta}\rangle_{t-\tau(\alpha_t)}\|_2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \\
&\leq 4\gamma_{\max}b_{\max}\mathbf{E}[6A_{\max}\|\langle\tilde{\theta}\rangle_t\|_2 + 5(b_{\max}+\mu_{\max}) \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\leq 12\gamma_{\max}A_{\max}b_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2 \mid \mathcal{F}_{t-\tau(\alpha_t)}] \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k \\
&\quad + (12A_{\max}+20b_{\max}+20\mu_{\max})\gamma_{\max}b_{\max} \sum_{k=t-\tau(\alpha_t)}^{t-1} \alpha_k. \qquad (105)
\end{aligned}$$

Then, by using equation 100–equation 105, we have

$$
\begin{aligned}
&|\mathbf{E}[(\langle\tilde{\theta}\rangle_t-\theta^*)^\top(P+P^\top)(A(X_t)\langle\tilde{\theta}\rangle_t+B(X_t)^\top\pi_{t+1}-A\langle\tilde{\theta}\rangle_t-b)\mid\mathcal{F}_{t-\tau(\alpha_t)}]|\\
&\leq 54\alpha_t\gamma_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]+36\alpha_t\gamma_{\max}\frac{(b_{\max}+\mu_{\max})^2}{A_{\max}^2}+2\alpha_t\gamma_{\max}\|\theta^*\|_2^2\\
&\quad+2\gamma_{\max}\alpha_t\left(1+9\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]+6\frac{(b_{\max}+\mu_{\max})^2}{A_{\max}^2}+\|\theta^*\|_2^2\right)\\
&\quad+216\gamma_{\max}A_{\max}^2\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k\\
&\quad+152\gamma_{\max}\left(b_{\max}+\mu_{\max}+A_{\max}\|\theta^*\|_2\right)^2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k\\
&\quad+72\,\gamma_{\max}A_{\max}(2A_{\max}+b_{\max}+\mu_{\max})\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k\\
&\quad+48\gamma_{\max}A_{\max}(b_{\max}+\mu_{\max})(\frac{b_{\max}+\mu_{\max}}{A_{\max}}+1)^2\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k\\
&\quad+96A_{\max}^2\gamma_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k+67(b_{\max}+\mu_{\max})^2\gamma_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k\\
&\quad+12\gamma_{\max}A_{\max}b_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k\\
&\quad+(12A_{\max}+20b_{\max}+20\mu_{\max})\gamma_{\max}b_{\max}\sum_{k=t-\tau(\alpha_t)}^{t-1}\alpha_k\\
&\leq\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\left(72+456A_{\max}^2+84A_{\max}b_{\max}+72A_{\max}\mu_{\max}\right)\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2\mid\mathcal{F}_{t-\tau(\alpha_t)}]\\
&\quad+\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\Bigg[2+4\|\theta^*\|_2^2+48\frac{(b_{\max}+\mu_{\max})^2}{A_{\max}^2}+152\left(b_{\max}+\mu_{\max}+A_{\max}\|\theta^*\|_2\right)^2\\
&\quad+12A_{\max}b_{\max}+48A_{\max}(b_{\max}+\mu_{\max})(\frac{b_{\max}+\mu_{\max}}{A_{\max}}+1)^2+87(b_{\max}+\mu_{\max})^2\Bigg],
\end{aligned}
$$

where we use $\alpha_t\leq\alpha_{t-\tau\alpha_t}$ from Assumption 5 and $\tau(\alpha_t)\geq 1$ in the last inequality. This completes the proof. ■

Lemma 24 *Suppose that Assumptions 2–4 hold and $\alpha_t=\frac{\alpha_0}{t+1}$. When $\mu_t+\tau(\alpha_t)\alpha_{t-\tau(\alpha_t)}\zeta_8\leq\frac{0.1}{\gamma_{\max}}$ and $\tau(\alpha_t)\alpha_{t-\tau(\alpha_t)}\leq\min\{\frac{\log 2}{A_{\max}},\ \frac{0.1}{\zeta_8\gamma_{\max}}\}$, we have for $t\geq\bar{T}$,*

$$
\begin{aligned}
\mathbf{E}[\|\langle\tilde{\theta}\rangle_{t+1}-\theta^*\|_2^2]\leq\frac{\bar{T}}{t+1}\frac{\gamma_{\max}}{\gamma_{\min}}\mathbf{E}[\|\langle\tilde{\theta}\rangle_{\bar{T}}-\theta^*\|_2^2]+\frac{\zeta_9\alpha_0C\log^2(\frac{t+1}{\alpha_0})}{t+1}\frac{\gamma_{\max}}{\gamma_{\min}}\\
+\alpha_0\frac{\gamma_{\max}}{\gamma_{\min}}\frac{\sum_{l=\bar{T}}^{t+1}\mu_l}{t+1},
\end{aligned}
$$

where $\bar{T}$ is defined in Appendix A.2, ζ_8 and ζ_9 are defined in equation 18 and equation 19, respectively.

Proof of Lemma 24: Let $H(\langle\tilde{\theta}\rangle_t)=(\langle\tilde{\theta}\rangle_t-\theta^*)^\top P(\langle\tilde{\theta}\rangle_t-\theta^*)$. From Assumption 4, we know that

$$\gamma_{\min}\|\langle\tilde{\theta}\rangle_t-\theta^*\|_2^2\leq H(\langle\tilde{\theta}\rangle_t)\leq\gamma_{\max}\|\langle\tilde{\theta}\rangle_t-\theta^*\|_2^2.$$

Recall the update of $\langle\tilde{\theta}\rangle_t$ in equation 90:

$$\langle\tilde{\theta}\rangle_{t+1}=\langle\tilde{\theta}\rangle_t+\alpha_tA(X_t)\langle\tilde{\theta}\rangle_t+\frac{\alpha_t}{N}\sum_{i=1}^N b^i(X_t)+\alpha_t\rho_t.$$

From Assumption 2, for $t \geq \bar{T}$ we have

$$
\begin{aligned}
& H(\langle\tilde{\theta}\rangle_{t+1}) \\
&= (\langle\tilde{\theta}\rangle_{t+1} - \theta^*)^\top P(\langle\tilde{\theta}\rangle_{t+1} - \theta^*) \\
&= \left(\langle\tilde{\theta}\rangle_t + \alpha_t A(X_t)\langle\tilde{\theta}\rangle_t + \frac{\alpha_t}{N}\sum_{i=1}^N b^i(X_t) + \alpha_t\rho_t - \theta^*\right)^\top P \cdot \\
&\quad \left(\langle\tilde{\theta}\rangle_t + \alpha_t A(X_t)\langle\tilde{\theta}\rangle_t + \frac{\alpha_t}{N}\sum_{i=1}^N b^i(X_t) + \alpha_t\rho_t - \theta^*\right) \\
&= (\langle\tilde{\theta}\rangle_t - \theta^*)^\top P(\langle\tilde{\theta}\rangle_t - \theta^*) + \alpha_t^2(A(X_t)\langle\tilde{\theta}\rangle_t)^\top P(A(X_t)\langle\tilde{\theta}\rangle_t) \\
&\quad + \frac{\alpha_t^2}{N^2}(B(X_t)^\top \mathbf{1}_N)^\top P(B(X_t)^\top \mathbf{1}_N) + \frac{\alpha_t^2}{N}(A(X_t)\langle\tilde{\theta}\rangle_t)^\top (P + P^\top)(B(X_t)^\top \mathbf{1}_N) + \alpha_t^2\rho_t^\top P\rho_t \\
&\quad + \alpha_t^2(A(X_t)\langle\tilde{\theta}\rangle_t + \frac{1}{N}B(X_t)^\top \mathbf{1}_N)^\top (P + P^\top)\rho_t + \alpha_t(\langle\tilde{\theta}\rangle_t - \theta^*)^\top (P + P^\top)\rho_t \\
&\quad + \alpha_t(\langle\tilde{\theta}\rangle_t - \theta^*)^\top (P + P^\top)(A(X_t)\langle\tilde{\theta}\rangle_t + \frac{1}{N}B(X_t)^\top \mathbf{1}_N - A\langle\tilde{\theta}\rangle_t - b) \\
&\quad + \alpha_t(\langle\tilde{\theta}\rangle_t - \theta^*)^\top P(A\langle\tilde{\theta}\rangle_t + b) + \alpha_t(A\langle\tilde{\theta}\rangle_t + b)^\top P(\langle\tilde{\theta}\rangle_t - \theta^*) \\
&= H(\langle\tilde{\theta}\rangle_t) + \alpha_t^2(A(X_t)\langle\tilde{\theta}\rangle_t)^\top P(A(X_t)\langle\tilde{\theta}\rangle_t) + \frac{\alpha_t^2}{N^2}(B(X_t)^\top \mathbf{1}_N)^\top P(B(X_t)^\top \mathbf{1}_N) \\
&\quad + \frac{\alpha_t^2}{N}(A(X_t)\langle\tilde{\theta}\rangle_t)^\top (P + P^\top)(B(X_t)^\top \mathbf{1}_N) + \alpha_t^2\rho_t^\top P\rho_t \\
&\quad + \alpha_t^2(A(X_t)\langle\tilde{\theta}\rangle_t + \frac{1}{N}B(X_t)^\top \mathbf{1}_N)^\top (P + P^\top)\rho_t + \alpha_t(\langle\tilde{\theta}\rangle_t - \theta^*)^\top (P + P^\top)\rho_t \\
&\quad + \alpha_t(\langle\tilde{\theta}\rangle_t - \theta^*)^\top (P + P^\top)(A(X_t)\langle\tilde{\theta}\rangle_t + \frac{1}{N}B(X_t)^\top \mathbf{1}_N - A\langle\tilde{\theta}\rangle_t - b) \\
&\quad + \alpha_t(\langle\tilde{\theta}\rangle_t - \theta^*)^\top (PA + A^\top P)(\langle\tilde{\theta}\rangle_t - \theta^*),
\end{aligned}
\tag{106}
$$

where we use the fact that $A\theta^* + b = 0$ on the last equality.

Next, we can take expectation on both sides of equation 106. From Assumption 4 and Lemma 23, for $t \geq \bar{T}$ we have

$$
\begin{aligned}
&\mathbf{E}[H(\langle\tilde{\theta}\rangle_{t+1})] \\
&= \mathbf{E}[H(\langle\tilde{\theta}\rangle_t)] + \alpha_t^2 \mathbf{E}[(A(X_t)\langle\tilde{\theta}\rangle_t)^\top P(A(X_t)\langle\tilde{\theta}\rangle_t)] - \alpha_t \mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \theta^*\|_2^2] + \mathbf{E}[\alpha_t^2 \rho_t^\top P \rho_t] \\
&\quad + \frac{\alpha_t^2}{N^2}\mathbf{E}[(B(X_t)^\top \mathbf{1}_N)^\top P(B(X_t)^\top \mathbf{1}_N)] + \frac{\alpha_t^2}{N}\mathbf{E}[(A(X_t)\langle\tilde{\theta}\rangle_t)^\top (P+P^\top)(B(X_t)^\top \mathbf{1}_N)] \\
&\quad + \alpha_t^2 \mathbf{E}[(A(X_t)\langle\tilde{\theta}\rangle_t + \frac{1}{N}B(X_t)^\top \mathbf{1}_N)^\top (P+P^\top)\rho_t] + \alpha_t \mathbf{E}[(\langle\tilde{\theta}\rangle_t - \theta^*)^\top (P+P^\top)\rho_t] \\
&\quad + \alpha_t \mathbf{E}[(\langle\tilde{\theta}\rangle_t - \theta^*)^\top (P+P^\top)(A(X_t)\langle\tilde{\theta}\rangle_t + \frac{1}{N}B(X_t)^\top \mathbf{1}_N - A\langle\tilde{\theta}\rangle_t - b)] \\
&\leq \mathbf{E}[H(\langle\tilde{\theta}\rangle_t)] + \alpha_t^2 A_{\max}^2 \gamma_{\max} \mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2] - \alpha_t \mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \theta^*\|_2^2] + 2\alpha_t \gamma_{\max}\|\rho_t\|_2 \mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \theta^*\|_2] \\
&\quad + \alpha_t^2 \gamma_{\max} b_{\max}^2 + 2\alpha_t^2 \gamma_{\max} A_{\max} b_{\max} \mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2] + \alpha_t^2 \gamma_{\max}\mu_{\max}^2 \\
&\quad + 2\alpha_t^2 \gamma_{\max}\mu_{\max}(A_{\max}\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2] + b_{\max}) \\
&\quad + \alpha_t \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\left(72 + 456A_{\max}^2 + 84A_{\max}b_{\max} + 72A_{\max}\mu_{\max}\right)\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2] \\
&\quad + \alpha_t \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\Bigg[2 + 48\frac{(b_{\max}+\mu_{\max})^2}{A_{\max}^2} + 152\left(b_{\max}+\mu_{\max}+A_{\max}\|\theta^*\|_2\right)^2 \\
&\quad + 4\|\theta^*\|_2^2 + 12A_{\max}b_{\max} + 48A_{\max}(b_{\max}+\mu_{\max})(\frac{b_{\max}+\mu_{\max}}{A_{\max}}+1)^2 + 87(b_{\max}+\mu_{\max})^2\Bigg] \\
&\leq \mathbf{E}[H(\langle\tilde{\theta}\rangle_t)] + (-\alpha_t + \alpha_t\gamma_{\max}\|\rho_t\|_2)\mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \theta^*\|_2^2] + \alpha_t\gamma_{\max}\|\rho_t\|_2 \\
&\quad + \alpha_t \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\left(72 + 458A_{\max}^2 + 84A_{\max}b_{\max} + 72A_{\max}\mu_{\max}\right)\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2] \\
&\quad + \alpha_t \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\Bigg[2 + 48\frac{(b_{\max}+\mu_{\max})^2}{A_{\max}^2} + 152\left(b_{\max}+\mu_{\max}+A_{\max}\|\theta^*\|_2\right)^2 \\
&\quad + 4\|\theta^*\|_2^2 + 12A_{\max}b_{\max} + 48A_{\max}(b_{\max}+\mu_{\max})(\frac{b_{\max}+\mu_{\max}}{A_{\max}}+1)^2 + 89(b_{\max}+\mu_{\max})^2\Bigg].
\end{aligned}
$$

Using the facts that $\mathbf{E}[\|\langle\tilde{\theta}\rangle_t\|_2^2] \leq 2\mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \theta^*\|_2^2] + 2\|\theta^*\|_2^2$ and $\gamma_{\min}\|\langle\tilde{\theta}\rangle_t - \theta^*\|_2^2 \leq H(\langle\tilde{\theta}\rangle_t) \leq \gamma_{\max}\|\langle\tilde{\theta}\rangle_t - \theta^*\|_2^2$, then

$$
\begin{aligned}
&\mathbf{E}[H(\langle\tilde{\theta}\rangle_{t+1})] \\
&\leq \mathbf{E}[H(\langle\tilde{\theta}\rangle_t)] + (-\alpha_t + \alpha_t\gamma_{\max}\mu_t)\mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \theta^*\|_2^2] + \alpha_t\gamma_{\max}\mu_t \\
&\quad + 2\alpha_t^2\gamma_{\max}(b_{\max}+\mu_{\max})^2 \\
&\quad + 2\alpha_t \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\left(72 + 458A_{\max}^2 + 84A_{\max}b_{\max} + 72A_{\max}\mu_{\max}\right)\mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \theta^*\|_2^2] \\
&\quad + 2\alpha_t \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\left(72 + 458A_{\max}^2 + 84A_{\max}b_{\max} + 72A_{\max}\mu_{\max}\right)\|\theta^*\|_2^2 \\
&\quad + \alpha_t \alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\Bigg[2 + 48\frac{(b_{\max}+\mu_{\max})^2}{A_{\max}^2} + 152\left(b_{\max}+\mu_{\max}+A_{\max}\|\theta^*\|_2\right)^2 \\
&\quad + 4\|\theta^*\|_2^2 + 12A_{\max}b_{\max} + 48A_{\max}(b_{\max}+\mu_{\max})(\frac{b_{\max}+\mu_{\max}}{A_{\max}}+1)^2 + 87(b_{\max}+\mu_{\max})^2\Bigg] \\
&\leq \mathbf{E}[H(\langle\tilde{\theta}\rangle_t)] + (-\alpha_t + \alpha_t\gamma_{\max}\mu_t + \alpha_t\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\zeta_8)\mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \theta^*\|_2^2] \\
&\quad + \alpha_t\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\zeta_9 + \alpha_t\gamma_{\max}\mu_t.
\end{aligned}
$$

Moreover, from $\alpha_t = \frac{\alpha_0}{t+1}$, $\alpha_0 \geq \frac{\gamma_{max}}{0.9}$ and the definition of $\bar{T}$, for all $t \geq \bar{T}$ we have

$$\begin{aligned}
\mathbf{E}[H(\langle\tilde{\theta}\rangle_{t+1})] &\leq (1 - \frac{0.9\alpha_t}{\gamma_{\max}})\mathbf{E}[H(\langle\tilde{\theta}\rangle_t)] + \alpha_t\alpha_{t-\tau(\alpha_t)}\tau(\alpha_t)\gamma_{\max}\zeta_9 + \alpha_t\gamma_{\max}\mu_t \\
&\leq \frac{t}{t+1}\mathbf{E}[H(\langle\tilde{\theta}\rangle_t)] + \alpha_0\gamma_{\max}\frac{\mu_t}{t+1} + \frac{\alpha_0^2 C\log(\frac{t+1}{\alpha_0})\gamma_{\max}\zeta_9}{(t+1)(t-\tau(\alpha_t)+1)} \\
&\leq \frac{\bar{T}}{t+1}\mathbf{E}[H(\langle\tilde{\theta}\rangle_{\bar{T}})] + \alpha_0\gamma_{\max}\sum_{l=\bar{T}}^{t}\frac{\mu_l}{l+1}\Pi_{u=l+1}^{t}\frac{u}{u+1} \\
&\quad + \alpha_0^2\gamma_{\max}\zeta_9\sum_{l=\bar{T}}^{t}\frac{C\log(\frac{l+1}{\alpha_0})}{(l+1)(l-\tau(\alpha_l)+1)}\Pi_{u=l+1}^{t}\frac{u}{u+1} \\
&= \frac{\bar{T}}{t+1}\mathbf{E}[H(\langle\tilde{\theta}\rangle_{\bar{T}})] + \alpha_0\gamma_{\max}\sum_{l=\bar{T}}^{t}\frac{\mu_l}{t+1} + \frac{\alpha_0^2\gamma_{\max}\zeta_9}{t+1}\sum_{l=\bar{T}}^{t}\frac{C\log(\frac{l+1}{\alpha_0})}{l-\tau(\alpha_l)+1} \\
&\leq \frac{\bar{T}}{t+1}\mathbf{E}[H(\langle\tilde{\theta}\rangle_{\bar{T}})] + \alpha_0\gamma_{\max}\frac{\sum_{l=\bar{T}}^{t}\mu_l}{t+1} + \frac{\alpha_0^2\gamma_{\max}\zeta_9}{t+1}\sum_{l=\bar{T}}^{t}\frac{2C\log(\frac{l+1}{\alpha_0})}{l+1} \\
&\leq \frac{\bar{T}}{t+1}\mathbf{E}[H(\langle\tilde{\theta}\rangle_{\bar{T}})] + \alpha_0\gamma_{\max}\frac{\sum_{l=\bar{T}}^{t+1}\mu_l}{t+1} + \frac{\zeta_9\alpha_0\gamma_{\max}C\log^2(\frac{t+1}{\alpha_0})}{t+1}, \qquad (107)
\end{aligned}$$

where we use

$$\sum_{l=\bar{T}}^{t}\frac{2\alpha_0\log(\frac{l+1}{\alpha_0})}{l+1} \leq \log^2(\frac{t+1}{\alpha_0})$$

to get the last inequality. Then, we can get the bound of $\mathbf{E}[\|\langle\tilde{\theta}\rangle_{t+1} - \theta^*\|_2^2]$ from equation 107 as follows:

$$\begin{aligned}
&\mathbf{E}[\|\langle\tilde{\theta}\rangle_{t+1} - \theta^*\|_2^2] \\
&\leq \frac{1}{\gamma_{\min}}\mathbf{E}[H(\langle\tilde{\theta}\rangle_{t+1})] \\
&\leq \frac{\bar{T}}{t+1}\frac{\gamma_{\max}}{\gamma_{\min}}\mathbf{E}[\|\langle\tilde{\theta}\rangle_{\bar{T}} - \theta^*\|_2^2] + \frac{\zeta_9\alpha_0 C\log^2(\frac{t+1}{\alpha_0})}{t+1}\frac{\gamma_{\max}}{\gamma_{\min}} + \alpha_0\frac{\gamma_{\max}}{\gamma_{\min}}\frac{\sum_{l=\bar{T}}^{t+1}\mu_l}{t+1}.
\end{aligned}$$

This completes the proof. ■

We are now in a position to prove Theorem 5.

Proof of Theorem 5: Note that

$$\sum_{i=1}^{N}\mathbf{E}[\|\theta_{t+1}^i - \theta^*\|_2^2] \leq 2\sum_{i=1}^{N}\mathbf{E}[\|\theta_{t+1}^i - \langle\tilde{\theta}\rangle_t\|_2^2] + 2N\mathbf{E}[\|\langle\tilde{\theta}\rangle_t - \theta^*\|_2^2].$$

By using Lemmas 20 and 24, for any $t \geq \bar{T}$, we have

$$\begin{aligned}
&\sum_{i=1}^{N}\mathbf{E}\left[\left\|\theta_{t+1}^i - \theta^*\right\|_2^2\right] \\
&\leq \frac{16}{\epsilon_1}\bar{\epsilon}^t\mathbf{E}[\|\sum_{i=1}^{N}\tilde{\theta}_0^i + \alpha_0 A(X_0)\theta_0^i + \alpha_0 b^i(X_0)\|_2] + \frac{16}{\epsilon_1}\frac{A_{\max}C_\theta + b_{\max}}{1-\bar{\epsilon}}\left(\alpha_0\bar{\epsilon}^{t/2} + \alpha_{\lceil\frac{t}{2}\rceil}\right) \\
&\quad + 2\alpha_t A_{\max}C_\theta + 2\alpha_t b_{\max} + \frac{2\bar{T}N}{t}\frac{\gamma_{\max}}{\gamma_{\min}}\mathbf{E}[\|\langle\tilde{\theta}\rangle_{\bar{T}} - \theta^*\|_2^2] \\
&\quad + \frac{2N\zeta_9\alpha_0 C\log^2(\frac{t}{\alpha_0})}{t}\frac{\gamma_{\max}}{\gamma_{\min}} + 2\alpha_0 N\frac{\gamma_{\max}}{\gamma_{\min}}\frac{\sum_{l=\bar{T}}^{t}\mu_l}{t} \\
&\leq C_7\bar{\epsilon}^t + C_8\left(\alpha_0\bar{\epsilon}^{\frac{t}{2}} + \alpha_{\lceil\frac{t}{2}\rceil}\right) + C_9\alpha_t + \frac{1}{t}\left(C_{10}\log^2\left(\frac{t}{\alpha_0}\right) + C_{11}\sum_{l=\bar{T}}^{t}\mu_l + C_{12}\right).
\end{aligned}$$

This completes the proof. ■

REFERENCES

Nicole Abaid and Maurizio Porfiri. Consensus over numerosity-constrained random networks. *IEEE Transactions on Automatic Control*, 56(3):649–654, 2010.

J. Bhandari, D. Russo, and R. Singal. A finite time analysis of temporal difference learning with linear function approximation. In *31st Conference on Learning Theory*, pp. 1691–1692, 2018.

V. S. Borkar. *Stochastic approximation: A Dynamical Systems Viewpoint*. Cambridge University Press, 2008.

V. S. Borkar and S. P. Meyn. The ODE method for convergence of stochastic approximation and reinforcement learning. *SIAM Journal on Control and Optimization*, 38(2):447–469, 2000.

V. S. Borkar and S. Pattathil. Concentration bounds for two time scale stochastic approximation. In *56th Annual Allerton Conference on Communication, Control, and Computing*, pp. 504–511, 2018.

S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.

P. Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*, volume 31. Springer Science & Business Media, 2013.

S. Chen, A. M. Devraj, A. Bušić, and S. Meyn. Explicit mean-square error bounds for monte-carlo and linear stochastic approximation. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 4173–4183. PMLR, 26–28 Aug 2020a.

Z. Chen, S. T. Maguluri, S. Shakkottai, and K. Shanmugam. Finite-sample analysis of contractive stochastic approximation using smooth convex envelopes. In *34th Conference on Neural Information Processing Systems*, 2020b.

G. Dalal, B. Szörényi, G. Thoppe, and S. Mannor. Finite sample analyses for TD(0) with function approximation. In *32nd AAAI Conference on Artificial Intelligence*, pp. 6144–6160, 2018a.

G. Dalal, G. Thoppe, B. Szörényi, and S. Mannor. Finite sample analysis of two-timescale stochastic approximation with applications to reinforcement learning. In *Conference On Learning Theory*, pp. 1199–1233. PMLR, 2018b.

P. Dayan. The convergence of TD (λ) for general λ. *Machine Learning*, 8(3-4):341–362, 1992.

T.T. Doan, S.T. Maguluri, and J. Romberg. Finite-time analysis of distributed TD(0) with linear function approximation on multi-agent reinforcement learning. In *36th International Conference on Machine Learning*, pp. 1626–1635, 2019.

T.T. Doan, S. T. Maguluri, and J. 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.

H. Gupta, R. Srikant, and L. Ying. Finite-time performance bounds and adaptive learning rate selection for two time-scale reinforcement learning. In *33rd Conference on Neural Information Processing System*, pp. 4706–4715, 2019.

J. Hajnal and M. S. Bartlett. Weak ergodicity in non-homogeneous Markov chains. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54:233–246, 1958.

H. J. Kushner. An averaging method for stochastic approximations with discontinuous dynamics, constraints, and state dependent noise. In *Recent advances in statistics*, pp. 211–235. Elsevier, 1983.

H.J. Kushner and G. Yin. Asymptotic properties of distributed and communicating stochastic approximation algorithms. *SIAM Journal on Control and Optimization*, 25(5):1266–1290, 1987.

C. Lakshminarayanan and C. Szepesvari. Linear stochastic approximation: How far does constant step-size and iterate averaging go? In *International Conference on Artificial Intelligence and Statistics*, pp. 1347–1355, 2018.

H. J. LeBlanc, H. Zhang, X. Koutsoukos, and S. Sundaram. Resilient asymptotic consensus in robust networks. *IEEE Journal on Selected Areas in Communications*, 31(4):766–781, 2013.

S. Ma, Y. Zhou, and S. Zou. Variance-reduced off-policy TDC learning: Non-asymptotic convergence analysis. In *34th Conference on Neural Information Processing Systems*, 2020.

A. S. Mathkar and V. S. Borkar. Nonlinear gossip. *SIAM Journal on Control and Optimization*, 54 (3):1535–1557, 2016.

A. Nedić and J. Liu. On convergence rate of weighted-averaging dynamics for consensus problems. *IEEE Transactions on Automatic Control*, 62(2):766–781, 2017.

A. Nedić and A. Olshevsky. Distributed optimization over time-varying directed graphs. *IEEE Transactions on Automatic Control*, 60(3):601–615, 2015.

G. Qu and A. Wierman. Finite-time analysis of asynchronous stochastic approximation and Q-learning. In *33rd Conference on Learning Theory*, volume 125, pp. 3185–3205. Proceedings of Machine Learning Research, 09–12 Jul 2020.

H. Robbins and S. Monro. A stochastic approximation method. *The annals of mathematical statistics*, pp. 400–407, 1951.

W. J. Rugh. *Linear System Theory (2nd Ed.)*. Prentice-Hall, Inc., USA, 1996. ISBN 0134412052.

R. Srikant and L. Ying. Finite-time error bounds for linear stochastic approximation and TD learning. In *32nd Conference on Learning Theory*, volume 99, pp. 2803–2830. Proceedings of Machine Learning Research, 25–28 Jun 2019.

R.S. Sutton and A.G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 2018.

B. Touri. *Product of Random Stochastic Matrices and Distributed Averaging*. Springer Science & Business Media, 2012.

J.N. Tsitsiklis and B. Van Roy. An analysis of temporal-difference learning with function approximation. *IEEE Transactions on Automatic Control*, 42(5):674–690, 1997.

N. H. Vaidya, L. Tseng, and G. Liang. Iterative approximate byzantine consensus in arbitrary directed graphs. In *Proceedings of the 2012 ACM symposium on Principles of distributed computing*, pp. 365–374, 2012.

G. Wang, B. Li, and G. B. Giannakis. A multistep lyapunov approach for finite-time analysis of biased stochastic approximation. *arXiv preprint arXiv:1909.04299*, 2019.

Y. Wang and S. Zou. Finite-sample analysis of Greedy-GQ with linear function approximation under markovian noise. In Jonas Peters and David Sontag (eds.), *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 124 of *Proceedings of Machine Learning Research*, pp. 11–20. PMLR, 03–06 Aug 2020.

Y. Wang, W. Chen, Y. Liu, Z. Ma, and T. Liu. Finite sample analysis of the GTD policy evaluation algorithms in markov setting. In *31st Conference on Neural Information Processing Systems*, pp. 5504–5513, 2017.

W. Weng, H. Gupta, N. He, L. Ying, and R. Srikant. The mean-squared error of double Q-learning. In *34th Conference on Neural Information Processing Systems*, 2020.

Y. Wu, W. Zhang, P. Xu, and Q. Gu. A finite time analysis of two time-scale actor critic methods. In *34th Conference on Neural Information Processing Systems*, 2020.

L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. In *Proceedings of the 4th International Conference on Information Processing in Sensor Networks*, pp. 63–70, 2005.

P. Xu and Q. Gu. A finite-time analysis of Q-learning with neural network function approximation. In *37th International Conference on Machine Learning*, 2020.

T. Xu, S. Zou, and Y. Liang. Two time-scale off-policy TD learning: Non-asymptotic analysis over markovian samples. In *33rd Conference on Neural Information Processing Systems*, pp. 10634–10644, 2019.

K. Zhang, Z. Yang, H. Liu, T. Zhang, and T. Başar. Fully decentralized multi-agent reinforcement learning with networked agents. In *35th International Conference on Machine Learning*, pp. 5872–5881, 2018.

S. Zou, T. Xu, and Y. Liang. Finite-sample analysis for SARSA with linear function approximation. In *33rd Conference on Neural Information Processing Systems*, pp. 8668–8678, 2019.