Supplementary materials for “Finite-Sample Analysis of Decentralized Temporal-Difference Learning with Linear Function Approximation”

A Proof of Theorem 1

Proof. From the definition of $\boldsymbol{G}(\boldsymbol{\Theta})$ in (24), we have that

$$\begin{aligned}\boldsymbol{G}(\boldsymbol{\Theta}(k),\xi_k) &= \begin{bmatrix}\boldsymbol{\theta}_1^\top(k)[\gamma\boldsymbol{\phi}(s(k+1))-\boldsymbol{\phi}(s(k))]\boldsymbol{\phi}^\top(s(k))\\ \boldsymbol{\theta}_2^\top(k)[\gamma\boldsymbol{\phi}(s(k+1))-\boldsymbol{\phi}(s(k))]\boldsymbol{\phi}^\top(s(k))\\ \vdots\\ \boldsymbol{\theta}_M^\top(k)[\gamma\boldsymbol{\phi}(s(k+1))-\boldsymbol{\phi}(s(k))]\boldsymbol{\phi}^\top(s(k))\end{bmatrix} + \begin{bmatrix} r_1(k)\boldsymbol{\phi}^\top(s(k))\\ r_2(k)\boldsymbol{\phi}^\top(s(k))\\ \vdots \\ r_M(k)\boldsymbol{\phi}^\top(s(k))\end{bmatrix}\\
&= \boldsymbol{\Theta}(k)\big[\gamma\boldsymbol{\phi}(s(k+1))-\boldsymbol{\phi}(s(k))\big]\boldsymbol{\phi}^\top(s(k)) + \boldsymbol{r}(k)\boldsymbol{\phi}^\top(s(k))\\
&= \boldsymbol{\Theta}(k)\boldsymbol{H}^\top(\xi_k) + \boldsymbol{r}(k)\boldsymbol{\phi}^\top(s(k))\end{aligned}$$

where we have used the definitions that $\boldsymbol{r}(k) = [r_1(k)\ r_2(k)\ \cdots\ r_M(k)]^\top$ and $\boldsymbol{H}(\xi_k) := \boldsymbol{\phi}(s(k))[\gamma\boldsymbol{\phi}^\top(s(k+1)) - \boldsymbol{\phi}^\top(s(k))]$. Using standard norm inequalties, it follows that

$$\begin{aligned}\|\boldsymbol{\Delta}\boldsymbol{G}(\boldsymbol{\Theta}(k),\xi_k)\|_F &\le \big\|[\gamma\boldsymbol{\phi}(s(k+1))-\boldsymbol{\phi}(s(k))]\,\boldsymbol{\phi}^\top(s(k))\big\|_F\cdot\|\boldsymbol{\Delta\Theta}(k)\|_F + \big\|\boldsymbol{r}(k)\boldsymbol{\phi}^\top(s(k))\big\|_F\\
&\le \big[\|\gamma\boldsymbol{\phi}(s(k+1))\|_F + \|\boldsymbol{\phi}(s(k))\|_F\big]\cdot\|\boldsymbol{\phi}^\top(s(k))\|_F\cdot\|\boldsymbol{\Delta\Theta}(k)\|_F + \|\boldsymbol{r}(k)\|_F\cdot\|\boldsymbol{\phi}(s(k))\|_F\\
&\le (1+\gamma)\|\boldsymbol{\Delta\Theta}(k)\|_F + \sqrt{M}r_{\max}\\
&\le 2\|\boldsymbol{\Delta\Theta}(k)\|_F + \sqrt{M}r_{\max}\end{aligned}\tag{44}$$

where $1+\gamma \le 2$ for the discounting factor $0 \le \gamma < 1$, and the last inequality holds since feature vectors $\|\boldsymbol{\phi}(s)\| \le 1$, rewards $r(k) \le r_{\max}$, and the Frobenious norm of rank-one matrices is equivalent to the ℓ_2-norm of vectors. For future reference, notice from the above inequality that $\lambda_{\max}(\boldsymbol{H}(\xi_k)) \le \|\boldsymbol{H}(\xi_k)\|_F = \big\|[\gamma\boldsymbol{\phi}(s(k+1)) - \boldsymbol{\phi}(s(k))]\,\boldsymbol{\phi}^\top(s(k))\big\| \le 1+\gamma \le 2$, for all $k \in \mathbb{N}^+$.

It follows from (28) that

$$\begin{aligned}\|\boldsymbol{\Delta\Theta}(k+1)\|_F &\le \|\boldsymbol{W}\boldsymbol{\Delta\Theta}(k)\|_F + \alpha\|\boldsymbol{\Delta}\boldsymbol{G}(\boldsymbol{\Theta}(k))\|_F\\
&\le \big[\lambda_2^{\boldsymbol{W}} + 2\alpha\big]\,\|\boldsymbol{\Delta\Theta}(k)\|_F + \alpha\sqrt{M}r_{\max}\end{aligned}\tag{45}$$

where the second inequality is obtained after using (44), and the following inequality [Nedić et al., 2018, Ma et al., 2019]

$$\|\boldsymbol{W}\boldsymbol{\Delta\Theta}(k)\|_F = \left\|\boldsymbol{W}\left(\boldsymbol{I} - \frac{1}{M}\mathbf{1}\mathbf{1}^\top\right)\boldsymbol{\Theta}(k)\right\| \le \lambda_2^{\boldsymbol{W}}\|\boldsymbol{\Delta\Theta}(k)\|_F.\tag{46}$$

Then applying (45) recursively from iteration k to 0 gives rise to

$$\begin{aligned}\|\boldsymbol{\Delta\Theta}(k)\|_F &\le \big(\lambda_2^{\boldsymbol{W}} + 2\alpha\big)^k\|\boldsymbol{\Delta\Theta}(0)\|_F + \alpha\sqrt{M}r_{\max}\sum_{i=0}^{k-1}\big(\lambda_2^{\boldsymbol{W}} + 2\alpha\big)^i\\
&\le \big(\lambda_2^{\boldsymbol{W}} + 2\alpha\big)^k\|\boldsymbol{\Delta\Theta}(0)\|_F + \frac{\alpha\sqrt{M}r_{\max}}{1-\lambda_2^{\boldsymbol{W}} - 2\alpha}\\
&\le \big(\lambda_2^{\boldsymbol{W}} + 2\alpha\big)^k\|\boldsymbol{\Delta\Theta}(0)\|_F + \alpha\cdot\frac{2\sqrt{M}r_{\max}}{1-\lambda_2^{\boldsymbol{W}}}\end{aligned}\tag{47}$$

where the last inequality is a consequence of using the fact that $0 < \alpha < \frac{1}{2}\cdot\frac{1-\lambda_2^{\boldsymbol{W}}}{2}$. This concludes the proof of Theorem 1. □

B Proof of Lemma 1

Proof. Recalling the definitions of $\boldsymbol{H}(\xi_k)$ ($\bar{\boldsymbol{H}}$) and $\boldsymbol{b}(\xi_k)$ ($\bar{\boldsymbol{b}}$), it is not difficult to verify that in the stationary distribution $\boldsymbol{\pi}$ of the Markov chain, the expectations of $\boldsymbol{H}(\xi_k)$ and $\boldsymbol{b}(\xi_k)$ obey

$$\mathbb{E}_\pi[\boldsymbol{H}(\xi_k)] = \bar{\boldsymbol{H}} \tag{48}$$

and

$$\mathbb{E}_\pi[\boldsymbol{b}_\mathcal{G}(\xi_k)] = \bar{\boldsymbol{b}}_\mathcal{G}. \tag{49}$$

Thus,

$$\mathbb{E}_\pi\left[\frac{1}{M}\boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_k)\mathbf{1}\Big|\mathcal{F}(k)\right] = \mathbb{E}_\pi\left[\boldsymbol{H}(\xi_k)\bar{\boldsymbol{\theta}}(k) + \boldsymbol{b}_\mathcal{G}(\xi_k)\big|\mathcal{F}(k)\right] = \bar{\boldsymbol{H}}\bar{\boldsymbol{\theta}}(k) + \bar{\boldsymbol{b}}_\mathcal{G} \tag{50}$$

and its variance satisfies

$$\begin{aligned}
\mathbb{E}_\pi\left[\left\|\frac{1}{M}\boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_k)\mathbf{1} - \bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}(k))\right\|^2\Big|\mathcal{F}(k)\right] &= \mathbb{E}_\pi\left[\left\|(\boldsymbol{H}(\xi_k) - \bar{\boldsymbol{H}})\bar{\boldsymbol{\theta}}(k) + \boldsymbol{b}_\mathcal{G}(\xi_k) - \bar{\boldsymbol{b}}_\mathcal{G}\right\|^2\big|\mathcal{F}(k)\right] \\
&\le \mathbb{E}_\pi\left[2\left\|(\boldsymbol{H}(\xi_k) - \bar{\boldsymbol{H}})\bar{\boldsymbol{\theta}}(k)\right\|^2 + 2\left\|\boldsymbol{b}_\mathcal{G}(\xi_k) - \bar{\boldsymbol{b}}_\mathcal{G}\right\|^2\big|\mathcal{F}(k)\right] \\
&\le 2\beta^2\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^* + \boldsymbol{\theta}^*\|^2 + 8r_{\max}^2 \\
&\le 4\beta^2\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2 + 4\beta^2\|\boldsymbol{\theta}^*\|^2 + 8r_{\max}^2
\end{aligned} \tag{51}$$

where β denotes the largest absolute value of eigenvalues of $\boldsymbol{H}(\xi_k) - \bar{\boldsymbol{H}}$, for any $k \in \mathbb{N}^+$. □

C Proof of Theorem 2

Proof. Clearly, it holds that

$$\begin{aligned}
\mathbb{E}_\pi[\|\bar{\boldsymbol{\theta}}(k+1) - \boldsymbol{\theta}^*\|^2\big|\mathcal{F}(k)] &= \mathbb{E}_\pi\left[\left\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^* + \alpha\frac{1}{M}\boldsymbol{G}^\top(\boldsymbol{\Theta},\xi_k)\mathbf{1}\right\|^2\Big|\mathcal{F}(k)\right] \\
&\le \|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2 + 2\alpha\left\langle \bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*, \mathbb{E}_\pi\left[\frac{1}{M}\boldsymbol{G}(\boldsymbol{\Theta}(k),\xi_k)^T\mathbf{1}\Big|\mathcal{F}(k)\right]\right\rangle \\
&\quad + \alpha^2\mathbb{E}_\pi\left[\left\|\frac{1}{M}\boldsymbol{G}(\boldsymbol{\Theta}(k),\xi_k)^T\mathbf{1} - \bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}(k)) + \bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}(k))\right\|^2\Big|\mathcal{F}(k)\right] \\
&\le \|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2 + 2\alpha\left\langle \bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}(k)) - \bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}^*)\right\rangle \\
&\quad + 2\alpha^2(\beta^2\|\bar{\boldsymbol{\theta}}\|^2 + r_{\max}^2) + 2\alpha^2\|\bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}(k)) - \bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}^*)\|^2 \\
&\le \|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2 + 2\alpha\left\langle \bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*, \bar{\boldsymbol{H}}(\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*)\right\rangle \\
&\quad + 2\alpha^2(4\beta^2\|\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\|^2 + 4\beta^2\|\boldsymbol{\theta}^*\|^2 + 8r_{\max}^2) + 2\alpha^2\|\bar{\boldsymbol{H}}(\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*)\|^2 \\
&\le \left[1 + 2\alpha\lambda_{\max}^{\bar{\boldsymbol{H}}} + 8\alpha^2\beta^2 + 2\alpha^2(\lambda_{\min}^{\bar{\boldsymbol{H}}})^2\right]\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2 \\
&\quad + (8\alpha^2\beta^2\|\boldsymbol{\theta}^*\|^2 + 16\alpha^2 r_{\max}^2).
\end{aligned} \tag{52}$$

where $\lambda_{\max}^{\bar{\boldsymbol{H}}}$ and $\lambda_{\min}^{\bar{\boldsymbol{H}}}$ are the largest and the smallest eigenvalues of $\bar{\boldsymbol{H}}$, respectively. Because $\bar{\boldsymbol{H}}$ is a negative definite matrix, then it follows that $\lambda_{\min}^{\bar{\boldsymbol{H}}} < \lambda_{\max}^{\bar{\boldsymbol{H}}} < 0$.

Defining constants $c_1 := 1 + 2\alpha\lambda_{\max}^{\bar{\boldsymbol{H}}} + 8\alpha^2\beta^2 + 2\alpha^2(\lambda_{\min}^{\bar{\boldsymbol{H}}})^2$, and choosing any constant stepsize α obeying $0 < \alpha \le -\frac{1}{2}\cdot\frac{\lambda_{\max}^{\bar{\boldsymbol{H}}}}{4\beta^2 + (\lambda_{\min}^{\bar{\boldsymbol{H}}})^2}$, then we have $c_1 < 1$ and $\frac{1}{1-c_1} \le -\frac{1}{\alpha\lambda_{\max}^{\bar{\boldsymbol{H}}}}$. Now, taking expectation with respect to $\mathcal{F}(k)$ in (52) gives rise to

$$\mathbb{E}\left[\|\bar{\boldsymbol{\theta}}(k+1) - \boldsymbol{\theta}^*\|^2\right] \le c_1\mathbb{E}\left[\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2\right] + (8\alpha^2\beta^2\|\boldsymbol{\theta}^*\|^2 + 16\alpha^2 r_{\max}^2). \tag{53}$$

Applying the above recursion from iteration k to iteration 0 yields

$$\mathbb{E}\left[\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2\right] \le c_1^k\|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 + \frac{1-c_1^k}{1-c_1}\left(8\alpha^2\beta^2\|\boldsymbol{\theta}^*\|^2 + 16\alpha^2 r_{\max}^2\right)$$

$$\le c_1^k \|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 + \frac{8\alpha^2\beta^2\|\boldsymbol{\theta}^*\|^2 + 16\alpha^2 r_{\max}^2}{-\alpha\lambda_{\max}^{\bar{\boldsymbol{H}}}}$$
$$\le c_1^k \|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 + \alpha c_2 \tag{54}$$

where $c_2 := \frac{8\beta^2\|\boldsymbol{\theta}^*\|^2+16r_{\max}^2}{-\lambda_{\max}^{\bar{\boldsymbol{H}}}}$, and this concludes the proof. □

D Proof of Proposition 1

Proof. We have that

$$\begin{aligned}
\mathbb{E}\big[\|\boldsymbol{\theta}_m(k) - \boldsymbol{\theta}^*\|^2\big] &= \mathbb{E}\big[\|\boldsymbol{\theta}_m(k) - \bar{\boldsymbol{\theta}}(k) + \bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2\big] \\
&\le 2\mathbb{E}\big[\|\boldsymbol{\theta}_m(k) - \bar{\boldsymbol{\theta}}(k)\|^2\big] + 2\mathbb{E}\big[\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2\big] \\
&\le 2\mathbb{E}\big[\|\boldsymbol{\Delta\Theta}(k)\|_F^2\big] + 2\mathbb{E}\big[\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2\big] \\
&\le 2\mathbb{E}\left[\big(\lambda_2^{\boldsymbol{W}} + 2\alpha\big)^k \|\boldsymbol{\Delta\Theta}(0)\|_F + \frac{2\alpha\sqrt{M} r_{\max}}{1-\lambda_2^{\boldsymbol{W}}}\right]^2 + 2c_1^k\|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 + 2\alpha c_2 \\
&\le 4\big(\lambda_2^{\boldsymbol{W}} + 2\alpha\big)^{2k}\|\boldsymbol{\Delta\Theta}(0)\|_F^2 + \frac{8\alpha^2 M r_{\max}^2}{(1-\lambda_2^{\boldsymbol{W}})^2} + 2c_1^k\|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 + 2\alpha c_2.
\end{aligned} \tag{55}$$

where the third inequality follows from using (29) and (54). Letting $c_3 := \max\{\big(\lambda_2^{\boldsymbol{W}} + 2\alpha\big)^2, c_1\}$, $V_0 := 2\max\{4\|\boldsymbol{\Delta\Theta}(0)\|_F^2, 2\|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2\}$, and $c_4 := \alpha \cdot \frac{8Mr_{\max}^2}{(1-\lambda_2^{\boldsymbol{W}})^2} + \frac{16\beta^2\|\boldsymbol{\theta}^*\|^2+32r_{\max}^2}{-\lambda_{\max}^{\bar{\boldsymbol{H}}}}$, then it is straightforward from (55) that our desired result follows; that is,

$$\mathbb{E}\Big[\|\boldsymbol{\theta}_m(k) - \boldsymbol{\theta}^*\|^2\Big] \le c_3^k V_0 + c_4\alpha \tag{56}$$

which concludes the proof. □

E Proof of Lemma 2

Proof. For notational brevity, let $r_{\mathcal{G}}(k) := (1/M)\sum_{m\in\mathcal{M}} r_m(k)$ for each $k \in \mathbb{N}^+$. It then follows that

$$\begin{aligned}
&\Big\|\frac{1}{KM}\sum_{j=k}^{k+K-1}\mathbb{E}\big[\boldsymbol{G}^\top(\boldsymbol{\Theta},\xi_j)\mathbf{1}\big|\mathcal{F}(k)\big] - \bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}})\Big\| \\
&= \Big\|\frac{1}{K}\sum_{j=k}^{k+K-1}\mathbb{E}\Big[\boldsymbol{\phi}(s(k))[\gamma\boldsymbol{\phi}(s(k+1)) - \boldsymbol{\phi}(s(k))]^\top\bar{\boldsymbol{\theta}} + \frac{1}{M}\boldsymbol{\phi}(s(k))\boldsymbol{r}^\top(k)\mathbf{1}\Big] - \mathbb{E}_\pi\left[\boldsymbol{g}(\bar{\boldsymbol{\theta}})\right]\Big\| \\
&= \Big\|\frac{1}{K}\sum_{j=k}^{k+K-1}\sum_{s\in\mathcal{S}}\Big(\Pr\big[s(j)=s|\mathcal{F}(k)\big] - \pi(s)\Big)\Big[\boldsymbol{\phi}(s)\big(\gamma P(s,s')\boldsymbol{\phi}(s') - \boldsymbol{\phi}(s)\big)^\top(\bar{\boldsymbol{\theta}} + \boldsymbol{\theta}^*) + r_{\mathcal{G}}(s)\boldsymbol{\phi}(s)\Big]\Big\| \\
&\le \max_{s,s'}\Big\|\boldsymbol{\phi}(s)\big[\gamma P(s,s')\boldsymbol{\phi}(s') - \boldsymbol{\phi}(s)\big]^\top(\bar{\boldsymbol{\theta}} + \boldsymbol{\theta}^*) + r_{\mathcal{G}}(s)\boldsymbol{\phi}(s)\Big\| \\
&\quad \times \frac{1}{K}\sum_{j=k}^{k+K-1}\sum_{s\in\mathcal{S}}\Big|\Pr[s(j)=s|\mathcal{F}(k)] - \pi(s)\Big| \\
&\le (1+\gamma)\big(\|\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\| + 2\|\boldsymbol{\theta}^*\| + r_{\max}\big) \times \frac{1}{K}\sum_{j=k}^{k+K-1}\nu_0\rho^k\cdot\rho^{j-k} \\
&\le \frac{(1+\gamma)\nu_0\rho^k}{(1-\rho)K}(\|\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\| + 2\|\boldsymbol{\theta}^*\| + r_{\max}) \\
&\le \sigma_k(K)\big(\|\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\| + 1\big)
\end{aligned} \tag{57}$$

where $\sigma_k(K) = \frac{(1+\gamma)\nu_0\rho^k}{(1-\rho)K} \times \max\{2\|\boldsymbol{\theta}^*\| + r_{\max}, 1\}$, and the second inequality arises from the fact that any finite-state, irreducible, and aperiodic Markov chains converges geometrically fast (with some initial constant $\nu_0 > 0$ and rate $0 < \rho < 1$) to its unique stationary distribution [Levin and Peres, 2017, Thm. 4.9]. Thus, we conclude that Lemma 2 holds true with monotonically decreasing function $\sigma(K)$ of $K \in \mathbb{N}^+$ as defined above. □

F Proof of Lemma 3

Proof. Recalling the definition of our multi-step Lyapunov function, we obtain that

$$\mathbb{E}\big[\mathbb{V}(k+1) - \mathbb{V}(k)\big|\mathcal{F}(k)\big] = \mathbb{E}\big[\|\bar{\boldsymbol{\theta}}(k+K) - \boldsymbol{\theta}^*\|^2 - \|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2\big|\mathcal{F}(k)\big]\,. \tag{58}$$

Thus, we should next derive the bound of the right hand side of above equation. Following from iterate (27), we can write

$$\bar{\boldsymbol{\theta}}(k+K) = \bar{\boldsymbol{\theta}}(k) + \frac{\alpha}{M}\sum_{j=k}^{k+K-1} \boldsymbol{G}^\top(\boldsymbol{\Theta}(j), \xi_j)\mathbf{1}. \tag{59}$$

As a consequence (without particular statement, the expectation in the rest of this proof is taken with respect to the ξ_k to ξ_{k+K-1} conditioned on ξ_0 to ξ_{k-1}),

$$\begin{aligned}
&\mathbb{E}\Big[\|\bar{\boldsymbol{\theta}}(k+K) - \boldsymbol{\theta}^*\|^2\big|\mathcal{F}(k)\Big] = \mathbb{E}\Big[\Big\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^* + \frac{\alpha}{M}\sum_{j=k}^{k+K-1}\boldsymbol{G}^\top(\boldsymbol{\Theta}(j),\xi_j)\mathbf{1}\Big\|^2\Big|\mathcal{F}(k)\Big] \\
&= \mathbb{E}\Big[\Big\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^* + \frac{\alpha}{M}\sum_{j=k}^{k+K-1}\Big[\boldsymbol{G}^\top(\boldsymbol{\Theta}(j),\xi_j)\mathbf{1} - \boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1} + \boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1}\Big]\Big\|^2\Big|\mathcal{F}(k)\Big] \\
&= \|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2 \\
&\quad + 2\alpha\mathbb{E}\Big[\Big\langle\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*, K\bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}(k)) + \frac{1}{M}\sum_{j=k}^{k+K-1}\Big[\boldsymbol{G}^\top(\boldsymbol{\Theta}(j),\xi_j)\mathbf{1} - \boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1} + \boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1}\Big] - K\bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}(k))\Big\rangle\Big|\mathcal{F}(k)\Big] \\
&\quad + \alpha^2\mathbb{E}\Big[\Big\|\frac{1}{M}\sum_{j=k}^{k+K-1}\Big[\boldsymbol{G}^\top(\boldsymbol{\Theta}(j),\xi_j)\mathbf{1} - \boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1} + \boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1}\Big]\Big\|^2\Big|\mathcal{F}(k)\Big] \\
&= \|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2 + \underbrace{2\alpha\mathbb{E}\Big[\big\langle\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*, K\bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}(k)) - K\bar{\boldsymbol{g}}(\boldsymbol{\theta}^*)\big\rangle\Big|\mathcal{F}(k)\Big]}_{\text{the second term}} \\
&\quad + \underbrace{2\alpha\mathbb{E}\Big[\Big\langle\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*, \sum_{j=k}^{k+K-1}\frac{1}{M}\Big[\boldsymbol{G}^\top(\boldsymbol{\Theta}(j),\xi_j)\mathbf{1} - \boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1}\Big]\Big\rangle\Big|\mathcal{F}(k)\Big]}_{\text{the third term}} \\
&\quad + \underbrace{2\alpha\mathbb{E}\Big[\Big\langle\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*, \sum_{j=k}^{k+K-1}\frac{1}{M}\boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1} - K\bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}})\Big\rangle\Big|\mathcal{F}(k)\Big]}_{\text{the fourth term}} \\
&\quad + \underbrace{\alpha^2\mathbb{E}\Big[\Big\|\frac{1}{M}\sum_{j=k}^{k+K-1}\Big[\boldsymbol{G}^\top(\boldsymbol{\Theta}(j),\xi_j)\mathbf{1} - \boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1} + \boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1}\Big]\Big\|^2\Big|\mathcal{F}(k)\Big]}_{\text{the last term}}
\end{aligned} \tag{60}$$

where the second and the third equality result from adding and subtracting the same terms and the last equality holds since $\bar{g}(\boldsymbol{\theta}^*) = 0$. In the following, we will bound the four terms in the above equality.

1) **Bounding the second term.** As a direct result of the definition of $\bar{g}(\boldsymbol{\theta})$, we have that $\bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}) - \bar{\boldsymbol{g}}(\boldsymbol{\theta}^*) = \bar{\boldsymbol{H}}(\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)$. Therefore, it holds that

$$2\alpha\mathbb{E}\Big[\big\langle\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*, K\bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}(k)) - K\bar{\boldsymbol{g}}(\boldsymbol{\theta}^*)\big\rangle\Big|\mathcal{F}(k)\Big] = 2\alpha K\mathbb{E}\left[(\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*)^\top\bar{\boldsymbol{H}}(\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*)|\mathcal{F}(k)\right]$$

$$\leq 2\alpha K \lambda_{\max}^{\bar{\boldsymbol{H}}} \|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2 \tag{61}$$

where $\lambda_{\max}^{\bar{\boldsymbol{H}}}$ is the largest eigenvalue of $\bar{\boldsymbol{H}}$. Because $\bar{\boldsymbol{H}}$ is a negative definite matrix, it holds that $\lambda_{\max}^{\bar{\boldsymbol{H}}} < 0$.

2) **Bounding the third term.** Defining first $\boldsymbol{p}(k, \boldsymbol{\Theta}(k), K) := \sum_{j=k}^{k+K-1} \frac{1}{M}\Big[\boldsymbol{G}^\top(\boldsymbol{\Theta}(j), \xi_j)\mathbf{1} - \boldsymbol{G}^\top(\boldsymbol{\Theta}(k), \xi_j)\mathbf{1}\Big]$, then it follows that

$$\begin{aligned}
\boldsymbol{p}(k, \boldsymbol{\Theta}(k), K) = & \sum_{j=k}^{k+K-2} \frac{1}{M}\left[\boldsymbol{G}^\top(\boldsymbol{\Theta}(j), \xi_j)\mathbf{1} - \boldsymbol{G}^\top(\boldsymbol{\Theta}(k), \xi_j)\mathbf{1}\right] \\
& + \frac{1}{M}\Big[\boldsymbol{G}^\top(\boldsymbol{\Theta}(k+K-1), \xi_{k+K-1})\mathbf{1} - \boldsymbol{G}^\top(\boldsymbol{\Theta}(k), \xi_{k+K-1})\mathbf{1}\Big] \\
= & \ \boldsymbol{p}(k, \boldsymbol{\Theta}(k), K-1) + \frac{1}{M}\Big[\boldsymbol{G}^\top(\boldsymbol{\Theta}(k+K-1), \xi_{k+K-1})\mathbf{1} - \boldsymbol{G}^\top(\boldsymbol{\Theta}(k), \xi_{k+K-1})\mathbf{1}\Big] \\
= & \ \boldsymbol{p}(k, \boldsymbol{\Theta}(k), K-1) + \boldsymbol{H}(k+K-1)[\bar{\boldsymbol{\theta}}(k+K-1) - \bar{\boldsymbol{\theta}}(k)].
\end{aligned}$$

Recalling that 2 is the largest absolute value of eigenvalues of $\boldsymbol{H}(k)$ for any $k \in \mathbb{N}^+$ (which clearly exists and is bounded due to the bounded feature vectors $\boldsymbol{\phi}(s)$ for any $s \in \mathcal{S}$), the norm of $\boldsymbol{p}(k, \boldsymbol{\Theta}(k), K)$ can be bounded as follows

$$\begin{aligned}
\|\boldsymbol{p}(k, \boldsymbol{\Theta}(k), K)\| \leq & \ \|\boldsymbol{p}(k, \boldsymbol{\Theta}(k), K-1\| + 2\|\bar{\boldsymbol{\theta}}(k+K-1) - \bar{\boldsymbol{\theta}}(k)\| \\
= & \ \|\boldsymbol{p}(k, \boldsymbol{\Theta}(k), K-1)\| + 2\alpha \Bigg\| \sum_{j=k}^{k+K-2} \frac{1}{M}\Big[\boldsymbol{G}^\top(\boldsymbol{\Theta}(j), \xi_j)\mathbf{1} - \boldsymbol{G}^\top(\boldsymbol{\Theta}(k), \xi_j)\mathbf{1}\Big] \\
& + \sum_{j=k}^{k+K-2} \frac{1}{M}\boldsymbol{G}^\top(\boldsymbol{\Theta}(k), \xi_j)\mathbf{1} \Bigg\| \\
\leq & \ (1+2\alpha)\|\boldsymbol{p}(k, \boldsymbol{\Theta}(k), K-1)\| + 2\sum_{j=k}^{k+K-2} \alpha\|\boldsymbol{H}(j)\bar{\boldsymbol{\theta}}(k) + b_{\mathcal{G}}\| \\
\leq & \ (1+2\alpha)\|\boldsymbol{p}(k, \boldsymbol{\Theta}(k), K-1)\| + 4\alpha\Bigg(\sum_{j=k}^{k+K-2} \|\bar{\boldsymbol{\theta}}(k)\| + \frac{r_{\max}}{2}\Bigg)
\end{aligned}$$

where the last inequality follows from $\|\boldsymbol{H}(j)\bar{\boldsymbol{\theta}}(k)\| \leq 2\|\bar{\boldsymbol{\theta}}(k)\|$ for any $j \geq 0$. Following the above recursion, we can write

$$\begin{aligned}
\|\boldsymbol{p}(k, \boldsymbol{\Theta}(k), K)\| \leq & \ (1+2\alpha)^K \|\boldsymbol{p}(k, \boldsymbol{\Theta}(k), 0)\| + 4\alpha K\|\bar{\boldsymbol{\theta}}(k)\| \sum_{j=0}^{K-1}(1+2\alpha)^j(K-1-j) \\
\leq & \ 4\alpha(\|\bar{\boldsymbol{\theta}}(k)\| + \frac{r_{\max}}{2}) \sum_{j=0}^{K-1}(1+2\alpha)^j(K-1-j)
\end{aligned} \tag{62}$$

where the second inequality because $\|\boldsymbol{p}(k, \boldsymbol{\Theta}(k), 0)\| = 0$.

For any positive constant $x \neq 1$ and $K \in \mathbb{N}^+$, the following equality holds

$$\sum_{j=0}^{K-1} x^j(K-1-j) = \frac{x^K - Kx + K - 1}{(1-x)^2}. \tag{63}$$

Substituting $x = (1+2\alpha)$ into (63) along with plugging the result into (62) yields

$$\|\boldsymbol{p}(k, \boldsymbol{\Theta}(k), K)\| \leq \frac{(1+2\alpha)^K - 2K\alpha - 1}{\alpha} K\|\bar{\boldsymbol{\theta}}(k)\|. \tag{64}$$

According to the mid-value theorem, there exists some suitable constant $\delta \in [0, 1]$ such that the following holds true

$$(1+2\alpha)^K = 1 + 2K\alpha + \frac{1}{2}K(K-1)(1+\delta(2\alpha)^{K-2}(2\alpha)^2$$

$$\leq 1 + 2K\alpha + \frac{1}{2}K^2(1+2\alpha)^{K-2}(2\alpha)^2. \tag{65}$$

Thus, it is clear that

$$\frac{(1+2\alpha)^K - 2K\alpha - 1}{\alpha} \leq 2\alpha K^2(1+2\alpha)^{K-2}. \tag{66}$$

Upon plugging (66) into (64), it follows that

$$\begin{aligned}\|\boldsymbol{p}(k,\boldsymbol{\Theta}(k),K)\| &\leq 2\alpha K^2(1+2\alpha)^{K-2}(\|\bar{\boldsymbol{\theta}}(k)\| + \frac{r_{\max}}{2})\\ &\leq 2\alpha K^2(1+2\alpha)^{K-2}(\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\| + \|\boldsymbol{\theta}^*\| + \frac{r_{\max}}{2}).\end{aligned} \tag{67}$$

Now, we turn to the third term in (60)

$$\begin{aligned}&2\alpha\mathbb{E}\Big[\Big\langle \bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*, \sum_{j=k}^{k+K-1}\frac{1}{M}\Big[\boldsymbol{G}^\top(\boldsymbol{\Theta}(j),\xi_j)\mathbf{1} - \boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1}\Big]\Big\rangle \Big|\mathcal{F}(k)\Big]\\ &= 2\alpha\mathbb{E}\left[\langle \bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*, \boldsymbol{p}(k,\boldsymbol{\Theta}(k),K)\rangle\,|\mathcal{F}(k)\right]\\ &\leq 2\alpha\mathbb{E}\left[\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\|\cdot\|\boldsymbol{p}(k,\boldsymbol{\Theta}(k),K)\|\big|\mathcal{F}(k)\right]\\ &= 2\alpha\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\|\cdot\mathbb{E}\left[\|\boldsymbol{p}(k,\boldsymbol{\Theta}(k),K)\|\big|\mathcal{F}(k)\right]\\ &\leq 4\alpha^2K^2(1+2\alpha)^{K-2}\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\|\cdot(\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\| + \|\boldsymbol{\theta}^*\| + \frac{r_{\max}}{2})\\ &\leq 4\alpha^2K^2(1+2\alpha)^{K-2}\Big(2\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\|^2 + \frac{1}{4}\|\boldsymbol{\theta}^*\|^2 + \frac{r_{\max}}{8}\Big).\end{aligned} \tag{68}$$

where the second inequality is obtained by plugging in (67), and the last one follows from the inequality $a(a+b) \leq 2a^2 + (1/4)b^2$.

3) Bounding the fourth term. It follows that

$$\begin{aligned}&2\alpha\mathbb{E}\Big[\Big\langle \bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*, \sum_{j=k}^{k+K-1}\frac{1}{M}\boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1} - K\bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}(k))\Big\rangle\Big|\mathcal{F}(k)\Big]\\ &= 2\alpha\Big\langle \bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*, \mathbb{E}\Big[\sum_{j=k}^{k+K-1}\frac{1}{M}\boldsymbol{G}(\boldsymbol{\Theta}(k),\xi_j))^T\mathbf{1} - K\bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}(k))\Big|\mathcal{F}(k)\Big]\Big\rangle\\ &\leq 2\alpha\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\|\cdot\Big\|\mathbb{E}\Big[\sum_{j=k}^{k+K-1}\frac{1}{M}\boldsymbol{G}(\boldsymbol{\Theta}(k),\xi_j))^T\mathbf{1} - K\bar{\boldsymbol{g}}(\bar{\boldsymbol{\theta}}(k))\Big|\mathcal{F}(k)\Big]\Big\|\\ &\leq 2\alpha K\sigma(K)\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\|(\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\|+1)\\ &\leq 2\alpha K\sigma(K)\Big(2\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\|^2 + \frac{1}{4}\Big).\end{aligned} \tag{69}$$

4) Bounding the last term. Evidently, we have that

$$\begin{aligned}&\Big\|\frac{1}{M}\sum_{j=k}^{k+K-1}\Big[\boldsymbol{G}^\top(\boldsymbol{\Theta}(j),\xi_j)\mathbf{1} - \boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1} + \boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1}\Big]\Big\|^2\\ &\leq 2\left\|\boldsymbol{p}(k,\boldsymbol{\Theta}(k),K)\right\|^2 + 2\Big\|\sum_{j=k}^{k+K-1}\frac{1}{M}\boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1}\Big\|^2\\ &\leq 2\left\|\boldsymbol{p}(k,\boldsymbol{\Theta}(k),K)\right\|^2 + 2\Big\|\sum_{j=k}^{k+K-1}\boldsymbol{H}(j)\bar{\boldsymbol{\theta}}(k) + \frac{1}{M}\boldsymbol{r}^\top(j)\mathbf{1}\boldsymbol{\phi}(j)\Big\|^2\\ &\leq 16\alpha^2K^4(1+2\alpha)^{2K-4}\|\bar{\boldsymbol{\theta}}(k)\|^2 + 16K\|\bar{\boldsymbol{\theta}}(k)\|^2 + \left[\alpha^2K^4(1+2\alpha)^{2K-4} + 4K\right]r_{\max}^2\end{aligned}$$

$$\leq \Big[32\alpha^2 K^6(1+2\alpha)^{2K-4}+32K\Big]\big(\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\|^2+\|\boldsymbol{\theta}^*\|^2\big)+\big[\alpha^2K^4(1+2\alpha)^{2K-4}+4K\big]r_{\max}^2 \tag{70}$$

where the first and the last inequality is the result of $\|\sum_{i=1}^n \boldsymbol{x}_i\|^2 \leq n\sum_{i=1}^n \|\boldsymbol{x}_i\|^2$ for any $\boldsymbol{x}$ and n; and the second is obtained by plugging in (67). Hence, upon taking expectation of both sides of (70) conditioning on $\mathcal{F}(k)$, we arrive at

$$\alpha^2\mathbb{E}\bigg[\bigg\|\frac{1}{M}\sum_{j=k}^{k+K-1}\Big[\boldsymbol{G}^\top(\boldsymbol{\Theta}(j),\xi_j)\mathbf{1}-\boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1}+\boldsymbol{G}^\top(\boldsymbol{\Theta}(k),\xi_j)\mathbf{1}\Big]\bigg\|^2\Big|\mathcal{F}(k)\bigg]$$
$$\leq\big[32\alpha^4K^6(1+2\alpha)^{2K-4}+32K\alpha^2\big]\big(\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\|^2+\|\boldsymbol{\theta}^*\|^2\big)+\alpha^2\big[\alpha^2K^4(1+2\alpha)^{2K-4}+4K\big]r_{\max}^2. \tag{71}$$

We have successfully bounded each of the four terms in (60). Putting now together the bounds in (61), (68), (69), and (71) into (60), we finally arrive at

$$\mathbb{E}\Big[\big\|\bar{\boldsymbol{\theta}}(k+K)-\boldsymbol{\theta}^*\big\|^2\big|\mathcal{F}(k)\Big]\leq\big[1+2\alpha T\lambda_{\max}^{\bar{\boldsymbol{H}}}+\alpha\Gamma_1(\alpha,K)\big]\big\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\big\|^2+\alpha\Gamma_2(\alpha,K) \tag{72}$$

where

$$\Gamma_1(\alpha,K)=32\alpha^3K^4(1+2\alpha)^{2K-4}+32K\alpha+8\alpha K^2(1+2\alpha)^{K-2}+4K\sigma(K) \tag{73}$$
$$\Gamma_2(\alpha,K)=\big[32\alpha^3K^4(1+2\alpha)^{2K-4}+32K\alpha+\alpha K^2(1+2\alpha)^{K-2}\big]\|\boldsymbol{\theta}^*\|^2$$
$$+\big[4\alpha^3K^4(1+2\alpha)^{2K-4}+\frac{1}{2}\alpha K^2(1+2\alpha)^{K-2}+4\alpha K\big]r_{\max}^2+\frac{1}{2}K\sigma(K) \tag{74}$$

From the definition of our multi-step Lyapunov function, we obtain that

$$\begin{aligned}\mathbb{E}\big[\mathbb{V}(k+1)-\mathbb{V}(k)\big|\mathcal{F}(k)\big]&=\mathbb{E}\Big[\big\|\bar{\boldsymbol{\theta}}(k+K)-\boldsymbol{\theta}^*\big\|^2\big|\mathcal{F}(k)\Big]-\big\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\big\|^2\\&\leq\alpha[2K\lambda_{\max}^{\bar{\boldsymbol{H}}}+\Gamma_1(\alpha,K)]\big\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\big\|^2+\alpha\Gamma_2(\alpha,K)\\&\leq\alpha[2K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}}+\Gamma_1(\alpha_{\max},K_{\mathcal{G}})]\big\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\big\|^2+\alpha\Gamma_2(\alpha_{\max},K_{\mathcal{G}})\end{aligned} \tag{75}$$

where the last inequality is due to the fact that functions $\Gamma_1(\alpha,K_{\mathcal{G}})$ and $\Gamma_2(\alpha,K_{\mathcal{G}})$ are monotonically increasing in α. This concludes the proof. □

G Proof of Lemma 4

Proof. It is straightforward to check that

$$\begin{aligned}\big\|\bar{\boldsymbol{\theta}}(k+i)-\boldsymbol{\theta}^*\big\|^2&=\Big\|\bar{\boldsymbol{\theta}}(k+i-1)-\boldsymbol{\theta}^*+\frac{\alpha}{M}\boldsymbol{G}^\top(\boldsymbol{\Theta}(k+i-1),\xi_{k+i-1})\mathbf{1}\\&\quad-\frac{\alpha}{M}\boldsymbol{G}^\top(\mathbf{1}(\boldsymbol{\theta}^*)^\top,\xi_{k+i-1})\mathbf{1}+\frac{\alpha}{M}\boldsymbol{G}^\top(\mathbf{1}(\boldsymbol{\theta}^*)^\top,\xi_{k+i-1})\mathbf{1}\Big\|^2\\&\leq\|\bar{\boldsymbol{\theta}}(k+i-1)-\boldsymbol{\theta}^*\|^2]+3\alpha^2\|\boldsymbol{H}(k)(\bar{\boldsymbol{\theta}}(k+i-1)-\boldsymbol{\theta}^*)\|^2\\&\quad+3\alpha^2\Big\|\boldsymbol{H}(k)\boldsymbol{\theta}^*+\frac{1}{M}\boldsymbol{\phi}(s(k))\boldsymbol{r}^\top(k)\mathbf{1}\Big\|^2\\&\leq(3+12\alpha^2)\|\bar{\boldsymbol{\theta}}(k+i-1)-\boldsymbol{\theta}^*\|^2+6\alpha^2\big[4\|\boldsymbol{\theta}^*\|^2+r_{\max}^2\big]\\&\leq(3+12\alpha^2)^i\|\bar{\boldsymbol{\theta}}(k)-\boldsymbol{\theta}^*\|^2+6\alpha^2\big[4\|\boldsymbol{\theta}^*\|^2+r_{\max}^2\big]\sum_{j=0}^{i-1}(3+12\alpha^2)^j.\end{aligned} \tag{76}$$

As as result, $\mathbb{V}(k)$ can be bounded as

$$\begin{aligned}
\mathbb{V}(k) &= \sum_{i=0}^{K_{\mathcal{G}}-1} \|\bar{\boldsymbol{\theta}}(k+i) - \boldsymbol{\theta}^*\|^2 \\
&\leq \sum_{i=0}^{K_{\mathcal{G}}-1} (3+12\alpha^2)^i \|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2 + 6\alpha^2(4\|\boldsymbol{\theta}^*\|^2 + r_{\max}) \sum_{i=1}^{K_{\mathcal{G}}-1}\sum_{j=0}^{i-1}(3+12\alpha^2)^j \\
&= \frac{(3+12\alpha^2)^{K_{\mathcal{G}}}-1}{2+12\alpha^2}\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2 \\
&\quad + \alpha^2 \frac{6(3+12\alpha^2)\big[(3+12\alpha^2)^{K_{\mathcal{G}}-1}-1\big] - 6K_{\mathcal{G}} + 6}{(2+12\alpha^2)^2}\big[4\|\boldsymbol{\theta}^*\|^2 + r_{\max}^2\big]
\end{aligned} \tag{77}$$

With $c_5 := \frac{(3+12\alpha_{\max}^2)^{K_{\mathcal{G}}}-1}{2+3\alpha_{\max}^2}$ and $c_6 := \frac{6(3+12\alpha_{\max}^2)\big[(3+12\alpha_{\max}^2)^{K_{\mathcal{G}}-1}-1\big]-6K_{\mathcal{G}}+6}{2+12\alpha_{\max}^2}\big(4\|\boldsymbol{\theta}^*\|^2 + r_{\max}^2\big)$, we conclude that

$$\mathbb{V}(k) \leq c_5\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2 + \alpha^2 c_6. \tag{78}$$

□

H Proof of Theorem 3

Proof. The convergence of $\mathbb{E}\Big[\big\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\big\|^2\Big]$ is separately addressed in two phases:
1) The time instant $k < k_\alpha$, with $k_\alpha = \max\{k|\rho^k \geq \alpha\}$, namely, it holds that $\alpha\sigma(K) \leq \sigma_k(K) \leq \sigma(K)$ for any $k < k_\alpha$;
2) The time instant $k \geq k_\alpha$, i.e., it holds that $\sigma_k(K) \leq \alpha\sigma(K)$ for any $k \geq k_\alpha$.

Convergence of the first phase

From Lemma 4, we have

$$-\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2 \leq -\frac{1}{c_5}\mathbb{V}(k) + \frac{\alpha^2 c_6}{c_5}. \tag{79}$$

Substituting (79) into (75), and rearanging the terms give the recursion of Lyapunov function as follows

$$\begin{aligned}
\mathbb{E}\big[\mathbb{V}(k+1)\big|\mathcal{F}(k)\big] &\leq \Big\{1 + \frac{1}{c_5}\big[2\alpha K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}} + \alpha\Gamma_1(\alpha_{\max}, K_{\mathcal{G}})\big]\Big\}\mathbb{E}\big[\mathbb{V}(k)\big|\mathcal{F}(k)\big] \\
&\quad + \alpha\Big\{\Gamma_2(\alpha, K_{\mathcal{G}}) - \frac{\alpha^2 c_6}{c_5}\big[2K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}} + \Gamma_1(\alpha_{\max}, K_{\mathcal{G}})\big]\Big\} \\
&\leq c_7\mathbb{E}\big[\mathbb{V}(k)\big|\mathcal{F}(k)\big] + \alpha c_8
\end{aligned} \tag{80}$$

where $c_7 := 1 + \frac{1}{2c_5}\alpha_{\max}K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}} \in (0,1)$; constant $c_8 := \Gamma_2(\alpha_{\max}, K_{\mathcal{G}}) - \frac{\alpha_{\max}^2 c_6}{c_5}K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}} > 0$, and the last inequality holds true because of (41).

Deducing from (80), we obtain that

$$\begin{aligned}
\mathbb{E}[\mathbb{V}(k)] &\leq c_7^k\mathbb{V}(0) + \alpha c_8\frac{1-c_7^k}{1-c_7} \\
&= c_5c_7^k\|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 + \alpha^2 c_6 c_7^k + \alpha c_8\frac{1-c_7^k}{1-c_7} \\
&\leq c_5c_7^k\|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 + \alpha^2 c_6 + \frac{\alpha c_8}{1-c_7} \qquad (81) \\
&= c_5c_7^k\|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 + \alpha^2 c_6 - \frac{2c_5c_8}{K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}}} \qquad (82)
\end{aligned}$$

Recalling the definition of Lyapunov function, it is obvious that

$$\mathbb{E}\big[\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2\big] \leq \mathbb{E}[\mathbb{V}(k)] \leq c_5c_7^k\|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 + \alpha^2 c_6 - \frac{2c_5c_8}{K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}}} \tag{83}$$

which finishes the proof of the first phase.

Convergence of the second phase
Without repeating similar derivation, we directly have that the following holds for $\sigma_k(K) \leq \alpha\sigma(K)$:

$$\Gamma_1(\alpha, K) := 32\alpha^3 K^4(1+2\alpha)^{2K-4} + 32K\alpha + 8\alpha K^2(1+2\alpha)^{K-2} + 4K\alpha\sigma(K) \tag{84}$$

$$\begin{aligned}\Gamma_2(\alpha, K) := &\left[32\alpha^3 K^4(1+2\alpha)^{2K-4} + 32K\alpha + \alpha K^2(1+2\alpha)^{K-2}\right]\|\boldsymbol{\theta}^*\|^2 \\ &+ \left[4\alpha^3 K^4(1+2\alpha)^{2K-4} + \frac{1}{2}\alpha K^2(1+2\alpha)^{K-2} + 4\alpha K\right] r_{\max}^2 + \frac{1}{2}K\alpha\sigma(K).\end{aligned} \tag{85}$$

Subsequently, we have the following recursion of $\mathbb{V}(k)$ that is similar to but slightly different from (80).

$$\mathbb{E}\left[\mathbb{V}(k+1)\middle|\mathcal{F}(k)\right] \leq c_7\mathbb{E}\left[\mathbb{V}(k)\middle|\mathcal{F}(k)\right] + \alpha^2 c_8', \quad \forall k \geq k_\alpha \tag{86}$$

where $c_8' := \left[16\alpha_{\max}^2 K_{\mathcal{G}}^6(1+2\alpha_{\max})^{2K_{\mathcal{G}}-4} + 32K_{\mathcal{G}} + 2K_{\mathcal{G}}^3(1+2\alpha_{\max})^{K_{\mathcal{G}}-2}\right]\|\boldsymbol{\theta}^*\|^2 + 4K_{\mathcal{G}} r_{\max}^2 - \frac{1}{8}K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}} - \frac{\alpha_{\max}c_6}{c_5}K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}}$. It is easy to check that $c_8' \geq c_8$ due to the fact that $\alpha_{\max} < 1$ in our case.
Repeatedly applying the above recursion from $k = k_\alpha$ to any $k > k_\alpha$ yields

$$\begin{aligned}\mathbb{E}[\mathbb{V}(k)] &\leq c_7^{k-k_\alpha}\mathbb{E}\left[\mathbb{V}(k_\alpha)\right] + \alpha^2 c_8' \frac{1-c_7^{k-k_\alpha}}{1-c_7} \\ &\leq c_7^{k-k_\alpha}\left(c_5 c_7^{k_\alpha}\|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 + \alpha^2 c_6 - \frac{2c_5c_8}{K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}}}\right) - \alpha\frac{2c_5c_8'}{K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}}} \\ &\leq c_5 c_7^k\|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 + c_7^{k-k_\alpha}\alpha^2 c_6 - (c_7^{k-k_\alpha} + \alpha)\frac{2c_5c_8'}{K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}}}\end{aligned} \tag{87}$$

where we have used $c_8 \leq c_8'$ for simplicity.

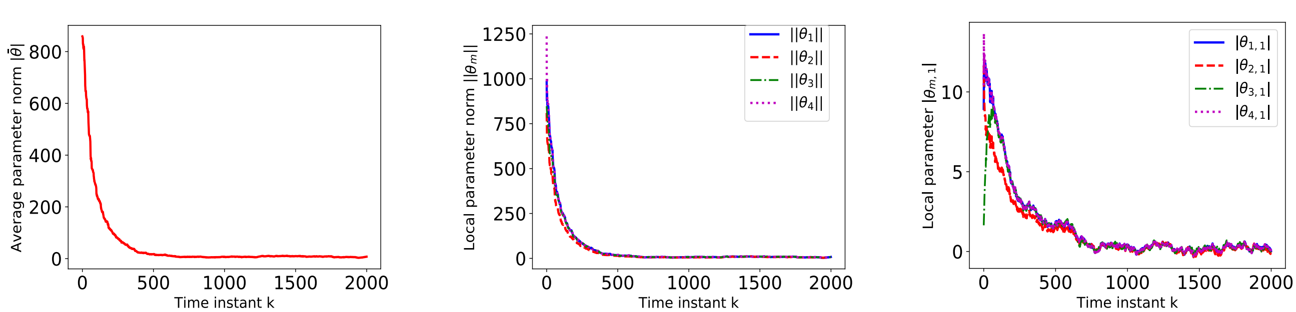
Again, using the definition of the Lyapunov function and (87), it follows that

$$\mathbb{E}\left[\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2\right] \leq c_5 c_7^k\|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 + c_7^{k-k_\alpha}\alpha^2 c_6 - (c_7^{k-k_\alpha} + \alpha)\frac{2c_5c_8'}{K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}}}, \quad \forall k \geq k_\alpha \tag{88}$$

Combining the results in the above two phases, we conclude that the following bound holds for any $k \in \mathbb{N}^+$

$$\mathbb{E}\left[\|\bar{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}^*\|^2\right] \leq c_5 c_7^k\|\bar{\boldsymbol{\theta}}(0) - \boldsymbol{\theta}^*\|^2 - \frac{2c_5c_8'}{K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}}}\alpha + \min\{1,\, c_7^{k-k_\alpha}\} \times \left(\alpha^2 c_6 - \frac{2c_5c_8'}{K_{\mathcal{G}}\lambda_{\max}^{\bar{\boldsymbol{H}}}}\right). \tag{89}$$

□



(a) Average parameter norm (b) Local parameters' norm (c) Local parameters

Figure 1: Consensus and convergence of decentralized TD(0) learning

I SIMULATIONS

In order to verify our analytical results, we carried out experiments on a multi-agent networked system. The details of our experimental setup are as follows: the number of agents $M = 30$, the state space size $|\mathcal{S}| = 100$ with each state s being a vector of length $|s| = 20$, the dimension of learning parameter $\boldsymbol{\theta}$ is $p = 10$, the reward upper bound $r^{\max} = 10$, and the stepsize $\alpha = 0.01$. The feature vectors are cosine functions, that is, $\boldsymbol{\phi}(s) = \cos(\boldsymbol{A}s)$, where $\boldsymbol{A} \in \mathcal{R}^{p\times|s|}$ is a randomly generated matrix. The communication weight matrix $\boldsymbol{W}$ depicting the neighborhood of the agents including the topology and the weights was generated randomly, with each agent being associated with 5 neighbors on average. As illustrated in Fig. 1(a), the parameter average $\bar{\boldsymbol{\theta}}$ converges to a small neighborhood of the optimum at a linear rate. To demonstrate the consensus among agents, convergence of the parameter norms $\|\boldsymbol{\theta}_m\|$ for $m = 1, 2, 3, 4$ is presented in Fig. 1(b), while that of their first elements $|\theta_{m,1}|$ is depicted in Fig. 1(c). The simulation results corroborate our theoretical analysis.