Supplementary Materials

I. Additional Materials

A. Explicit Form of Several Terms

In Theorem IV.3, the explicit forms of the terms are as follows,

$$\begin{aligned}
&\Upsilon(C(\mathbf{K})) \coloneqq 1 + \|\mathbf{B}\|\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})}\sqrt{\frac{\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|(C(\mathbf{K})-C(\mathbf{K}^*))}{\mu}}\\
&\mathbb{L} \coloneqq \left(2\sigma_n(\mathbf{R}) + \frac{2\|\mathbf{B}\|^2 C(\mathbf{K}(0))}{\sigma_1(\mathbf{\Psi})} + 4\sqrt{2}\xi\|\mathbf{B}\|\frac{C(\mathbf{K}(0))}{\sigma_1(\mathbf{\Psi})}\right)\frac{C(\mathbf{K}(0))}{\sigma_n(\mathbf{Q})}\\
&\xi \coloneqq \frac{1}{\sigma_n(\mathbf{Q})}\left[\frac{1+\|\mathbf{B}\|^2}{\sigma_1(\mathbf{\Psi})}C(\mathbf{K}(0)) + \sigma_1(\mathbf{R}) - 1\right]\\
&\varpi_2 \coloneqq -f_1^3(C(\mathbf{K}))\rho^{3(\kappa+1)}\\
&\varpi_1 \coloneqq -f_1^2(C(\mathbf{K}))\rho^{2(\kappa+1)} + \frac{\mathbb{L}}{2}\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}}C(\mathbf{K}))\|\\
&\varpi_0 \coloneqq \|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}}C(\mathbf{K}))\| - f_1^1(C(\mathbf{K}))\rho^{\kappa+1},
\end{aligned}$$

the function $f_1^1(\cdot)$, $f_1^2(\cdot)$ and $f_1^3(\cdot)$ are defined as,

$$\begin{aligned}
f_1^1(C(\mathbf{K})) \coloneqq& \frac{2\sqrt{d}C(\mathbf{K})}{\sigma_1(\mathbf{Q})}\mathcal{E}(C(\mathbf{K}))(\mathbb{C}\sum_i L_i)\\
f_1^2(C(\mathbf{K})) \coloneqq& 8\sqrt{d}(\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})})^2\frac{\|\mathbf{B}\|(\Upsilon(\mathbf{K})+1)}{\sigma_1(\mathbf{\Psi})}\mathcal{E}(C(\mathbf{K}))(\mathbb{C}\sum_i L_i)^2 + \frac{\sqrt{d}C(\mathbf{K})}{\sigma_1(\mathbf{Q})}\mathcal{E}(C(\mathbf{K}))(\|\mathbf{R}\| + \|\mathbf{B}\|^2\frac{C(\mathbf{K})}{\sigma_1(\mathbf{\Psi})})(\mathbb{C}\sum_i L_i)^2\\
f_1^3(C(\mathbf{K})) \coloneqq& 4\sqrt{d}(\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})})^2\frac{\|\mathbf{B}\|(\Upsilon(\mathbf{K})+1)}{\sigma_1(\mathbf{\Psi})}(\|\mathbf{R}\| + \|\mathbf{B}\|^2\frac{C(\mathbf{K})}{\sigma_1(\mathbf{\Psi})})(\eta\mathbb{C}\sum_i L_i)^3.
\end{aligned}$$

In inequality (IV.7), $M_1F_1(C(\mathbf{K}))$ and $M_2F_2(C(\mathbf{K}))$ are defined as,

$$M_1F_1(C(\mathbf{K}))\rho^{\kappa+1} \coloneqq \eta f_1^1(\cdot)\rho^{\kappa+1} + \eta^2 f_1^2(\cdot)\rho^{2(\kappa+1)} + \eta^3 f_1^3(\cdot)\rho^{3(\kappa+1)}$$

$$\begin{aligned}
M_2F_2(C(\mathbf{K}))\rho^r \coloneqq& \frac{2\eta C(\mathbf{K})}{\sigma_1(\mathbf{Q})}\mathcal{E}(C(\mathbf{K}))n\max_i(|\mathcal{N}_i^{-r}|)\rho^r + \eta^2\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})}\left[n\max_i(|\mathcal{N}_i^{-r}|)\rho^r\right]^2\left(\|\mathbf{R}\| + \|\mathbf{B}\|^2\frac{C(\mathbf{K})}{\mu}\right)\\
&+ 2\eta^2\frac{C(\mathbf{K})^2}{(\sigma_1(\mathbf{Q}))^2}\left[n\max_i(|\mathcal{N}_i^{-r}|)\rho^r\right]\left(\|\mathbf{R}\| + \|\mathbf{B}\|^2\frac{C(\mathbf{K})}{\mu}\right).
\end{aligned}$$

In Corollary IV.5, the bound of the second derivative $\mathbb{L}$ and the function $\mathcal{E}(C(\mathbf{K}))$ are defined as,

$$\mathbb{L} \coloneqq \left(2\sigma_n(\mathbf{R}) + \frac{2\|\mathbf{B}\|^2 C(\mathbf{K}(0))}{\sigma_1(\mathbf{\Psi})} + 4\sqrt{2}\zeta\|\mathbf{B}\|\frac{C(\mathbf{K}(0))}{\mu}\right)\frac{C(\mathbf{K}(0))}{\sigma_n(\mathbf{Q})}$$

$$\mathcal{E}(C(\mathbf{K})) \coloneqq \sqrt{\frac{(C(\mathbf{K})-C(\mathbf{K}^*))\left\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\frac{C(\mathbf{K})}{\sigma_1(\mathbf{\Psi})}\mathbf{B}\right\|}{\sigma_1(\mathbf{\Psi})}}$$

B. Helper Definitions and Lemmas

We recall the state transition dynamics from the global perspective:

$$\mathbf{x}(t+1) = (\mathbf{A}-\mathbf{B}\mathbf{K})\mathbf{x}(t) + \epsilon(\mathbf{t}),$$

where $\epsilon(\mathbf{t}) = \mathbf{w}(t) + \sigma_0\mathbf{B}\eta(t) \sim N(\mathbf{0}, \mathbf{\Psi})$ and we define $\mathbf{\Psi} = \mathbf{\Phi} + \sigma_0^2 \cdot \mathbf{B}\mathbf{B}^{\mathrm{T}}$ to simplify the notation.

For $\mathbf{K} \in \mathbb{R}^{d_{\mathbf{u}}\times d_{\mathbf{x}}}$ such that $\rho(\mathbf{A}-\mathbf{B}\mathbf{K}) < 1$, we define the operators:

$$\begin{aligned}
\mathcal{T}_{\mathbf{K}}(\mathbf{\Omega}) &= \sum_{t\geq 0}(\mathbf{A}-\mathbf{B}\mathbf{K})^t\mathbf{\Omega}[(\mathbf{A}-\mathbf{B}\mathbf{K})^t]^{\mathrm{T}},\\
\mathcal{T}_{\mathbf{K}}^{\mathrm{T}}(\mathbf{\Omega}) &= \sum_{t\geq 0}[(\mathbf{A}-\mathbf{B}\mathbf{K})^t]^{\mathrm{T}}\mathbf{\Omega}(\mathbf{A}-\mathbf{B}\mathbf{K})^t,
\end{aligned}$$

where $\mathbf{\Omega} \in \mathbb{R}^{d_\mathbf{x} \times d_\mathbf{x}}$ is a positive definite matrix. It is known that when $\rho(\mathbf{A}-\mathbf{BK})<1$, the Markov chain has stationary distribution $N(\mathbf{0}, \mathbf{\Xi_K})$, denoted by $\nu_\mathbf{K}$, where $\mathbf{\Xi_K}$ is the unique positive definite solution to the Lyapunov equation:

$$\mathbf{\Xi_K} = \mathbf{\Psi} + (\mathbf{A}-\mathbf{BK})\mathbf{\Xi_K}(\mathbf{A}-\mathbf{BK})^\mathrm{T}.$$

By definition, $\mathcal{T}_\mathbf{K}(\mathbf{\Omega})$ and $\mathcal{T}_\mathbf{K}^\mathrm{T}(\mathbf{\Omega})$ also satisfy Lyapunov equations:

$$\mathcal{T}_\mathbf{K}(\mathbf{\Omega}) = \mathbf{\Omega} + (\mathbf{A}-\mathbf{BK})\mathcal{T}_\mathbf{K}(\mathbf{\Omega})(\mathbf{A}-\mathbf{BK})^\mathrm{T},$$
$$\mathcal{T}_\mathbf{K}^\mathrm{T}(\mathbf{\Omega}) = \mathbf{\Omega} + (\mathbf{A}-\mathbf{BK})^\mathrm{T}\mathcal{T}_\mathbf{K}^\mathrm{T}(\mathbf{\Omega})(\mathbf{A}-\mathbf{BK}).$$

We have $\mathbf{\Xi_K} = \mathcal{T}_\mathbf{K}(\mathbf{\Psi})$.

Definition 1. [1] [Spatially exponential decaying (SED)] A matrix $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2}$ is (c, γ)-spatially exponential decaying (SED) if,

$$\|[\mathbf{X}]_{ij}\| \le c \cdot \gamma^{\mathrm{dist}(i,j)}, \quad \forall i, j \in \mathcal{N},$$

where $0<\gamma<1$, $c>0$.

Definition 2. [Local Spatially exponential decaying (L-SED)] A matrix $\mathbf{X}_i \in \mathbb{R}^{n_1 \times n_2}$ related to agent i is (c_i, γ_i)-local spatially exponential decaying (SED) if,

$$\|[\mathbf{X}_i]_{mn}\| \le c_i {\gamma_i}^{\mathrm{dist}(m,i)+\mathrm{dist}(i,n)},$$

where $0<\gamma_i<1$, $c_i>0$.

Definition 3. $i \to n_1 \to n_2 \to \ldots \to n_{t-1} \to j$ is defined as a walk of length t from i to j.

Definition 4. For convenience, we define an expanded connection graph $\mathcal{G}(r) = (\mathcal{N}, \mathcal{E}(r))$ based on the underlying graph $\mathcal{G}$. In $\mathcal{G}(r)$, agent i is connected to all its r-hop neighbors defined in $\mathcal{G}$.

Definition 5. We define $\mathcal{W}_{i\to j}^t(r)$ as a set that contains all the walks from i to j with length t, in a defined graph $\mathcal{E}(r)$.

Lemma 1. [Theorem 5.6.12 and Corollary 5.6.13 in [2]] There always exist a constant $H>0$ and $\alpha \in (0,1)$ such that $\|\mathbf{A}^t\| \le H\alpha^t$ if and only if $\rho(\mathbf{A})<1$ ($\mathbf{A}$ is stable).

Lemma 2. Suppose $\mathbf{X}$ and $\mathbf{Y}$ are square matrices of the same dimension, and they are (x,γ)-SED and (y,γ)-SED respectively, then the matrix $\mathbf{X}+\mathbf{Y}$ is $(x+y,\gamma)$-SED and the matrix $\mathbf{X}\cdot\mathbf{Y}$ is (Nxy,γ)-SED.

Proof.

$$\|[\mathbf{XY}]_{ij}\| = \|\sum_{k=1}^{n}[\mathbf{X}]_{ik}[\mathbf{Y}]_{kj}\| \le \sum_{k=1}^{n}\|[\mathbf{X}]_{ik}\|\|[\mathbf{Y}]_{kj}\| \le \sum_{k=1}^{n} xye^{-\gamma(\mathrm{dist}(i,k)+\mathrm{dist}(k,j))} \le \sum_{k=1}^{n} xye^{-\gamma\mathrm{dist}(i,j)} \le Nxye^{-\gamma\mathrm{dist}(i,j)}.$$

□

Lemma 3. Suppose $\mathbf{X}$ is (x,γ)-SED, $\mathbf{Y} \in \mathcal{M}^\kappa$ and $\max_{ij}\|[\mathbf{Y}]_{ij}\| = \bar{y}$, the $\mathbf{XY}$ is $(nx\bar{y}e^{\gamma\kappa},\gamma)$-SED, $\mathbf{X}+\mathbf{Y}$ is $(x+\frac{\widehat{y}}{e^{-\gamma\kappa}},\gamma)$-SED.

Proof.

$$\|[\mathbf{XY}]_{ij}\| = \|\sum_{k=1}^{n}[\mathbf{X}]_{ik}[\mathbf{Y}]_{kj}\| \le \sum_{k\in\mathcal{N}_j^\kappa}\|[\mathbf{X}]_{ik}\|\|[\mathbf{Y}]_{kj}\| \le \sum_{k\in\mathcal{N}_j^\kappa} x\bar{y}e^{-\gamma\mathrm{dist}(i,k)} \le \sum_{k\in\mathcal{N}_j^\kappa} x\bar{y}e^{-\gamma(\mathrm{dist}(i,j)-\kappa)} \le nx\bar{y}e^{\gamma\kappa}e^{-\gamma\mathrm{dist}(i,j)},$$

if $\mathrm{dist}(i,j) \le \kappa$:

$$\|[\mathbf{X}+\mathbf{Y}]_{ij}\| \le \|[\mathbf{X}]_{ij}\| + \|[\mathbf{Y}]_{ij}\| \le xe^{-\gamma\mathrm{dist}(i,j)} + \widehat{y} \le (x+\frac{\widehat{y}}{e^{-\gamma\mathrm{dist}(i,j)}})\cdot e^{-\gamma\mathrm{dist}(i,j)} \le (x+\frac{\widehat{y}}{e^{-\gamma\kappa}})\cdot e^{-\gamma\mathrm{dist}(i,j)},$$

else:

$$\|[\mathbf{X}+\mathbf{Y}]_{ij}\| = \|[\mathbf{X}]_{ij}\| \le xe^{-\gamma\mathrm{dist}(i,j)}.$$

□

Lemma 4. Suppose $\mathbf{X} \in \mathcal{M}^{\kappa_x}$ and $\max_{ij}\|[\mathbf{X}]_{ij}\| = \bar{x}$, $\mathbf{Y} \in \mathcal{M}^{\kappa_y}$ and $\max_{ij}\|[\mathbf{Y}]_{ij}\| = \bar{y}$, then $\mathbf{XY} \in \mathcal{M}^{\kappa_x+\kappa_y}$.

Proof.

$$[\mathbf{XY}]_{ij} = \sum_{k=1}^{n}[\mathbf{X}]_{ik}[\mathbf{Y}]_{kj} = \sum_{k\in\mathcal{N}_i^{\kappa_x}\cap\mathcal{N}_j^{\kappa_y}}[\mathbf{X}]_{ik}[\mathbf{Y}]_{kj},$$

if $\mathrm{dist}(i,j) > \kappa_x+\kappa_y$, then $[\mathbf{XY}]_{ij} = \mathbf{0}$.

□

$$\eta < \min \left\{ \begin{array}{l} \underbrace{\frac{1}{16}(\frac{\sigma_1(\mathbf{Q})\mu}{C(\mathbf{K})})^2 \frac{1}{\|\mathbf{B}\| \|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\|(1+\|\mathbf{A}-\mathbf{B}\mathbf{K}\|)}}_{\textit{To bound } \|\mathbf{\Xi}_{\mathbf{K}'} - \mathbf{\Xi}_{\mathbf{K}}\|}, \underbrace{\frac{\sigma_1(\mathbf{Q})}{32 C(\mathbf{K})\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|}}_{\textit{To guarantee the convergence}}, \\ \underbrace{\frac{\sigma_1(\mathbf{Q})\mu}{4C(\mathbf{K})\|\mathbf{B}\|(\Upsilon(C(\mathbf{K}))+1)\mathbb{C}\sqrt{d}\sum_i L_i \rho^{\kappa+1}}}_{\textit{To guarantee that } \mathbf{K}'' \textit{is stabilizing}}, \underbrace{\frac{2}{\mathbb{L}}}_{\textit{To guarantee that } \mathbf{K}' \textit{is stabilizing}}, \underbrace{\frac{-\varpi_1 - \sqrt{\varpi_1{}^2 - 4\varpi_1\varpi_2}}{2\varpi_2}}_{C(\mathbf{K}'')\leq C(\mathbf{K})}, 1 \end{array} \right\} \tag{1}$$

II. Proofs of the Main Results

Theorem 1 (Full version of Theorem IV.3)**.** Assuming that $\mathbf{K}^*$ is the centralized optimal controller, which is not truncated, suppose that 1) $C(\mathbf{K}(0))$ is finite, 2) Lemma IV.16 or Lemma IV.17 holds, 3) we have $\mu := \sigma_1(\mathbf{\Psi}) > 0$, if all agents conduct the policy update in Algorithm 1, for an appropriate choice of the step-size η shown in (1), and an adequate communication range limit κ that,

$$\kappa > \frac{1}{-\log\rho} \log \frac{2\sqrt{d} C(\mathbf{K})\mathcal{E}(C(\mathbf{K}))\mathbb{C}\sum_i L_i}{\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}}\mathbb{C}(\mathbf{K}))\|^2 \sigma_1(\mathbf{Q})} - 1,$$

if we conduct the process for T steps that,

$$T \geq \frac{\|\mathbf{\Xi}_{\mathbf{K}^*}\|}{\eta\mu^2\sigma_1(\mathbf{R})} \log \frac{C(\mathbf{K}(0)) - C(\mathbf{K}^*)}{\epsilon},$$

the distributed gradient descent enjoys the following performance bound:

$$C(\mathbf{K}(T)) - C(\mathbf{K}^*) \leq \epsilon + \frac{\|\mathbf{\Xi}_{\mathbf{K}^*}\|}{\eta\mu^2\sigma_1(\mathbf{R})}[M_1 F_1(C(\mathbf{K}(0)))\rho^{\kappa+1} + M_2 F_2(C(\mathbf{K}(0)))\rho^r],$$

where,

$$\begin{aligned}
\Upsilon(C(\mathbf{K})) &= 1 + \|\mathbf{B}\| \frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})} \sqrt{\frac{\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|(C(\mathbf{K}) - C(\mathbf{K}^*))}{\mu}} \\
\mathbb{L} &= \left(2\sigma_n(\mathbf{R}) + \frac{2\|\mathbf{B}\|^2 C(\mathbf{K}(0))}{\sigma_1(\mathbf{\Psi})} + 4\sqrt{2}\xi\|\mathbf{B}\|\frac{C(\mathbf{K}(0))}{\sigma_1(\mathbf{\Psi})}\right)\frac{C(\mathbf{K}(0))}{\sigma_n(\mathbf{Q})} \\
\xi &= \frac{1}{\sigma_n(\mathbf{Q})}\left[\frac{1+\|\mathbf{B}\|^2}{\sigma_1(\mathbf{\Psi})} C(\mathbf{K}(0)) + \sigma_1(\mathbf{R}) - 1\right] \\
\varpi_2 &= -f_1^3(C(\mathbf{K}))\rho^{3(\kappa+1)} \\
\varpi_1 &= -f_1^2(C(\mathbf{K}))\rho^{2(\kappa+1)} + \frac{\mathbb{L}}{2}\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\| \\
\varpi_0 &= \|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\| - f_1^1(C(\mathbf{K}))\rho^{\kappa+1}.
\end{aligned}$$

M_1, M_2 are scalars and $f_1^1(\cdot)$, $f_1^2(\cdot)$, $f_1^3(\cdot)$, $F_1(\cdot)$, $F_2(\cdot)$ and $\mathcal{E}(\cdot)$ are monotonically increasing functions regarding $C(\mathbf{K})$. They are defined in the next few subsections.

A. Difference between $\mathbf{K}'$ and $\mathbf{K}''$

As the establishment of Lemma IV.16 and Lemma IV.17 leads to the exponential decay property and the localized gradient approximation in Lemma IV.2, we have,

$$\begin{aligned}
\|\mathbf{K}'' - \mathbf{K}'\| &= \eta\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K})) - \widehat{\mathbf{h}}(\mathbf{K})\| \leq \eta\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K})) - \widehat{\mathbf{h}}(\mathbf{K})\|_F = \eta\sum_{i=1}^n \|\nabla_{\mathbf{K}_i} C(\mathbf{K}) - \widehat{\mathbf{h}}_i(\mathbf{K})\|_F \\
&\leq \eta\sqrt{d}\sum_{i=1}^n \|\nabla_{\mathbf{K}_i} C(\mathbf{K}) - \widehat{\mathbf{h}}_i(\mathbf{K})\| \leq \eta\mathbb{C}\sqrt{d}\sum_i L_i\rho^{\kappa+1},
\end{aligned} \tag{2}$$

where we use $\widehat{\mathbf{h}}(\mathbf{K}) \in \mathbb{R}^{d_{\mathbf{u}} \times d_{\mathbf{x}}}$ to represent the concatenate of all $\widehat{\mathbf{h}}_i(\mathbf{K})$.

B. Stabilizability of $\mathbf{K}'$

As is shown in [3], for a fixed controller $\mathbf{K}(0)$, the sub-level set $S_{C(\mathbf{K}(0))} = \{\mathbf{K} : C(\mathbf{K}) < C(\mathbf{K}(0))\}$ is compact and if we can find a scalar $L > 0$ such that $\max_{K \in S_{C(\mathbf{K}(0))}} \|\nabla^2 C(\mathbf{K})\| = L$, then we can choose a constant step $\eta < \frac{2}{L}$ to guarantee that $C(\mathbf{K}') = C(\mathbf{K} - \eta \mathcal{P}_{\mathcal{M}^r}(\nabla C(\mathbf{K}))) \leq C(\mathbf{K})$. In [3], authors also gives an upper bound of L (Lemma 13), which we use $\mathbb{L}$ to denote,

$$\mathbb{L} = \left(2\sigma_n(\mathbf{R}) + \frac{2\|\mathbf{B}\|^2 C(\mathbf{K}(0))}{\sigma_1(\mathbf{\Psi})} + 4\sqrt{2}\xi\|\mathbf{B}\|\frac{C(\mathbf{K}(0))}{\sigma_1(\mathbf{\Psi})}\right)\frac{C(\mathbf{K}(0))}{\sigma_n(\mathbf{Q})}.$$

It results in the condition for η:

$$\eta \leq \frac{2}{\mathbb{L}}. \tag{3}$$

C. Stabilizability of $\mathbf{K}''$

If we can show that $\|\mathbf{K}'' - \mathbf{K}'\|$ is small enough with the appropriate choice of step size η, then $\mathbf{K}'$ is guaranteed to be stabilizing (Lemma 11). Combining the result in Section II-A and Lemma 11, it results in the condition for η:

$$\eta \leq \frac{\sigma_1(\mathbf{Q})\mu}{4C(\mathbf{K}')\|\mathbf{B}\|(\|\mathbf{A}-\mathbf{B}\mathbf{K}'\|+1)\mathbb{C}\sqrt{d}\sum_i L_i\rho^{\kappa+1}}, \tag{4}$$

$C(\mathbf{K}') \leq C(\mathbf{K})$ is guaranteed. Then we bound $\|\mathbf{A} - \mathbf{B}\mathbf{K}'\|$,

$$\begin{aligned}
&\|\mathbf{A}-\mathbf{B}\mathbf{K}'\| = \|\mathbf{A}-\mathbf{B}\mathbf{K}+\mathbf{B}(\mathbf{K}-\mathbf{K}')\| \leq \|\mathbf{A}-\mathbf{B}\mathbf{K}\| + \eta\|\mathbf{B}\|\|\mathcal{P}_{\mathcal{M}^r}(\nabla C(\mathbf{K}))\| \leq \|\mathbf{A}-\mathbf{B}\mathbf{K}\| + \eta d^r\|\mathbf{B}\|\|\nabla C(\mathbf{K})\| \\
&\leq \|\mathbf{A}\| + \|\mathbf{B}\|\|\mathbf{K}\| + \eta d^r\|\mathbf{B}\|\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})}\sqrt{\frac{\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|(C(\mathbf{K})-C(\mathbf{K}^*))}{\sigma_1(\mathbf{\Psi})}} \\
&\leq \|\mathbf{A}\| + d^r\|\mathbf{B}\|\frac{1}{\sigma_1(\mathbf{R})}\left(\sqrt{\frac{\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|(C(\mathbf{K})-C(\mathbf{K}^*))}{\sigma_1(\mathbf{\Psi})}} + \|\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{A}\|\right) \\
&\quad + \|\mathbf{B}\|\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})}\sqrt{\frac{\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|(C(\mathbf{K})-C(\mathbf{K}^*))}{\sigma_1(\mathbf{\Psi})}} \\
&:= \Upsilon(C(\mathbf{K})),
\end{aligned} \tag{5}$$

where $\|\mathbf{K}\|$, $\|\nabla_{\mathbf{K}} C(\mathbf{K})\|$, $\|\mathbf{P}_{\mathbf{K}}\|$ can be bounded by the terms related to $C(\mathbf{K})$ as shown in Lemma 10 and Lemma 12 in the next section. d^r is a constant related to d and r because of the equivalence of norms (there exist positive constants c_1 and c_2 such that for all vectors $\mathbf{x}$ in the vector space: $c_1\|\mathbf{x}\| \leq \|\mathbf{x}\|_F \leq c_2\|\mathbf{x}\|$) and the fact that $\|\mathcal{P}_{\mathcal{M}^r}(\nabla C(\mathbf{K}))\|_F \leq \|\nabla C(\mathbf{K})\|_F$. Note that $\Upsilon(C(\mathbf{K}))$ monotonically increases with the increase of $C(\mathbf{K})$. We use $\eta \leq 1$ in the last inequality.

Based on (4) and the inequality above, we can newly introduce two conditions for η:

$$\eta \leq \min\left\{\frac{\sigma_1(\mathbf{Q})\mu}{4C(\mathbf{K})\|\mathbf{B}\|(\Upsilon(C(\mathbf{K}))+1)\mathbb{C}\sqrt{d}\sum_i L_i\rho^{\kappa+1}}, 1\right\}. \tag{6}$$

D. Descent guarantee of the objective function

1) Difference between $C(\mathbf{K})$ and $C(\mathbf{K}')$: Following the descent lemma in the $\mathbb{L}$-smooth function [4], if the step size η is smaller than $2/\mathbb{L}$, the one-step performance improvement by the exact gradient descent enjoys:

$$C(\mathbf{K}') - C(\mathbf{K}) \leq -(\eta - \frac{\mathbb{L}}{2}\eta^2)\left\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\right\|^2. \tag{7}$$

2) Difference between $C(\mathbf{K}')$ and $C(\mathbf{K}'')$: Note that $C(\mathbf{K}) = \mathbb{E}_{\mathbf{x}\sim N(\mathbf{0},\mathbf{\Psi})}\mathbf{x}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{x} + \sigma_0^2\operatorname{Tr}(\mathbf{R})$ (Lemma 6). For any $\mathbf{K}$, the noise remains the same and i.i.d. When we calculate the difference between two objective functions, the effect of noise can be removed. So we define a new state dynamics without noise, where $\mathbf{x}_t'' = (\mathbf{A}-\mathbf{B}\mathbf{K}'')^t\mathbf{x}$ for all $t \geq 0$ and $\mathbf{\Xi}_{\mathbf{K}''} = \mathbb{E}_{\mathbf{x}\sim N(\mathbf{0},\mathbf{\Psi})}[\sum_{t\geq 0}\mathbf{x}_t''(\mathbf{x}_t'')^{\mathrm{T}}]$. So that,

$$\begin{aligned}
C(\mathbf{K}'') - C(\mathbf{K}') =& \mathbb{E}_{\mathbf{x}\sim N(\mathbf{0},\mathbf{\Psi})}[\mathbf{x}^{\mathrm{T}}(\mathbf{P}_{\mathbf{K}''}-\mathbf{P}_{\mathbf{K}'})\mathbf{x}] = \mathbb{E}_{\mathbf{x}\sim N(\mathbf{0},\mathbf{\Psi})}\sum_{t\geq 0} A_{\mathbf{K}',\mathbf{K}''}(\mathbf{x}_t'') \\
=& 2\operatorname{Tr}\left[\mathbf{\Xi}_{\mathbf{K}''}(\mathbf{K}''-\mathbf{K}')\mathbf{E}_{\mathbf{K}}'\right] + \operatorname{Tr}\left[\mathbf{\Xi}_{\mathbf{K}''}(\mathbf{K}''-\mathbf{K}')^{\mathrm{T}}(\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}'\mathbf{B})(\mathbf{K}''-\mathbf{K}')\right] \\
\leq& 2\sqrt{d}\|\mathbf{\Xi}_{\mathbf{K}}''\|\|\mathbf{K}''-\mathbf{K}'\|\|\mathbf{E}_{\mathbf{K}}'\| + \sqrt{d}\|\mathbf{\Xi}_{\mathbf{K}}''\|(\|\mathbf{K}''-\mathbf{K}'\|)^2\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}'\mathbf{B}\| \\
\leq& 2\sqrt{d}\|\mathbf{\Xi}_{\mathbf{K}}''-\mathbf{\Xi}_{\mathbf{K}}'\|\|\mathbf{E}_{\mathbf{K}}'\|\|\mathbf{K}''-\mathbf{K}'\| + 2\sqrt{d}\|\mathbf{\Xi}_{\mathbf{K}}'\|\|\mathbf{E}_{\mathbf{K}}'\|\|\mathbf{K}''-\mathbf{K}'\| \\
&+ \sqrt{d}\|\mathbf{\Xi}_{\mathbf{K}}''-\mathbf{\Xi}_{\mathbf{K}}'\|\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}'\mathbf{B}\|(\|\mathbf{K}''-\mathbf{K}'\|)^2 + \sqrt{d}\|\mathbf{\Xi}_{\mathbf{K}}'\|\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}'\mathbf{B}\|(\|\mathbf{K}''-\mathbf{K}'\|)^2,
\end{aligned}$$

where $A_{\mathbf{K}',\mathbf{K}''}$ is the advantage function defined in Lemma 7. Here we avoid bounding $\|\Xi_{\mathbf{K}''}\|$ by $\frac{C(\mathbf{K}'')}{\sigma_1(\mathbf{Q})}$ (which is shown in Lemma 12) because we can not bound $C(\mathbf{K}'')$ by $C(\mathbf{K})$ so far.

We have three additional terms to bound: $\|\mathbf{E}_{\mathbf{K}'}\|$, $\|\Xi_{\mathbf{K}'}\|$, and $\|\Xi_{\mathbf{K}''}-\Xi_{\mathbf{K}'}\|$.

As $\mathbf{K}'$ is proved to be stabilizing, following Lemma 9, we have,

$$\begin{aligned}
\|\mathbf{E}_{\mathbf{K}'}\| &= \left\|(\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}'}\mathbf{B})\mathbf{K}'-\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}'}\mathbf{A}\right\| \\
&\leq \sqrt{\frac{(C(\mathbf{K}')-C(\mathbf{K}^*))\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}'}\mathbf{B}\|}{\sigma_1(\Psi)}} \\
&\leq \sqrt{\frac{(C(\mathbf{K}')-C(\mathbf{K}^*))\left\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\frac{C(\mathbf{K}')}{\sigma_1(\Psi)}\mathbf{B}\right\|}{\sigma_1(\Psi)}} \\
&\leq \sqrt{\frac{(C(\mathbf{K})-C(\mathbf{K}^*))\left\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\frac{C(\mathbf{K})}{\sigma_1(\Psi)}\mathbf{B}\right\|}{\sigma_1(\Psi)}} \\
&:= \mathcal{E}(C(\mathbf{K})),
\end{aligned} \tag{8}$$

where we use the bound of $\|\mathbf{P}_{\mathbf{K}}\|$ shown in Lemma 12 and the fact that $C(\mathbf{K}')\leq C(\mathbf{K})$. Note that $\mathcal{E}(C(\mathbf{K}))$ monotonically increases with the increase of $C(\mathbf{K})$.

Following Lemma 11 and conditioned on (6), (5), we have,

$$\|\Xi_{\mathbf{K}''}-\Xi_{\mathbf{K}'}\| \leq 4(\frac{C(\mathbf{K}')}{\sigma_1(\mathbf{Q})})^2\frac{\|\mathbf{B}\|(\|\mathbf{A}-\mathbf{B}\mathbf{K}'\|+1)}{\sigma_1(\Psi)}\|\mathbf{K}''-\mathbf{K}'\| \leq 4(\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})})^2\frac{\|\mathbf{B}\|(\Upsilon(\mathbf{K})+1)}{\sigma_1(\Psi)}\|\mathbf{K}''-\mathbf{K}'\|. \tag{9}$$

Following Lemma 12 and the fact that $C(\mathbf{K}')\leq C(\mathbf{K})$, we have,

$$\|\Xi_{\mathbf{K}'}\| \leq \frac{C(\mathbf{K}')}{\sigma_1(\mathbf{Q})} \leq \frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})}. \tag{10}$$

Combining (8), (9), (5), (10) and (2), we have that:

$$\begin{aligned}
&C(\mathbf{K}'')-C(\mathbf{K}') \\
\leq& 8\sqrt{d}(\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})})^2\frac{\|\mathbf{B}\|(\Upsilon(\mathbf{K})+1)}{\sigma_1(\Psi)}\mathcal{E}(C(\mathbf{K}))(\eta c\sum_i L_i\rho^{\kappa+1})^2 + \frac{2\sqrt{d}C(\mathbf{K})}{\sigma_1(\mathbf{Q})}\mathcal{E}(C(\mathbf{K}))(\eta c\sum_i L_i\rho^{\kappa+1}) \\
&+ 4\sqrt{d}(\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})})^2\frac{\|\mathbf{B}\|(\Upsilon(\mathbf{K})+1)}{\sigma_1(\Psi)}(\|\mathbf{R}\|+\|\mathbf{B}\|^2\frac{C(\mathbf{K})}{\sigma_1(\Psi)})(\eta c\sum_i L_i\rho^{\kappa+1})^3 \\
&+ \frac{\sqrt{d}C(\mathbf{K})}{\sigma_1(\mathbf{Q})}\mathcal{E}(C(\mathbf{K}))(\|\mathbf{R}\|+\|\mathbf{B}\|^2\frac{C(\mathbf{K})}{\sigma_1(\Psi)})(\eta c\sum_i L_i\rho^{\kappa+1})^2 \\
:=& \eta f_1^1(C(\mathbf{K}))\rho^{\kappa+1}+\eta^2 f_1^2(C(\mathbf{K}))\rho^{2(\kappa+1)}+\eta^3 f_1^3(C(\mathbf{K}))\rho^{3(\kappa+1)},
\end{aligned} \tag{11}$$

where we newly define three functions for notation simplification as follows,

$$\begin{aligned}
f_1^1(C(\mathbf{K})) =& \frac{2\sqrt{d}C(\mathbf{K})}{\sigma_1(\mathbf{Q})}\mathcal{E}(C(\mathbf{K}))(c\sum_i L_i) \\
f_1^2(C(\mathbf{K})) =& 8\sqrt{d}(\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})})^2\frac{\|\mathbf{B}\|(\Upsilon(\mathbf{K})+1)}{\sigma_1(\Psi)}\mathcal{E}(C(\mathbf{K}))(c\sum_i L_i)^2 + \frac{\sqrt{d}C(\mathbf{K})}{\sigma_1(\mathbf{Q})}\mathcal{E}(C(\mathbf{K}))(\|\mathbf{R}\|+\|\mathbf{B}\|^2\frac{C(\mathbf{K})}{\sigma_1(\Psi)})(c\sum_i L_i)^2 \\
f_1^3(C(\mathbf{K})) =& 4\sqrt{d}(\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})})^2\frac{\|\mathbf{B}\|(\Upsilon(\mathbf{K})+1)}{\sigma_1(\Psi)}(\|\mathbf{R}\|+\|\mathbf{B}\|^2\frac{C(\mathbf{K})}{\sigma_1(\Psi)})(\eta c\sum_i L_i)^3.
\end{aligned}$$

Note that these functions are also related to d, $\mathbf{Q}$, $\mathbf{B}$ and $\sigma_1(\Psi)$ except for $C(\mathbf{K})$. They are hidden for simplicity. Note that $f_1^1(C(\mathbf{K}))$, $f_1^2(C(\mathbf{K}))$, and $f_1^3(C(\mathbf{K}))$ monotonically increases with the increase of $C(\mathbf{K})$.

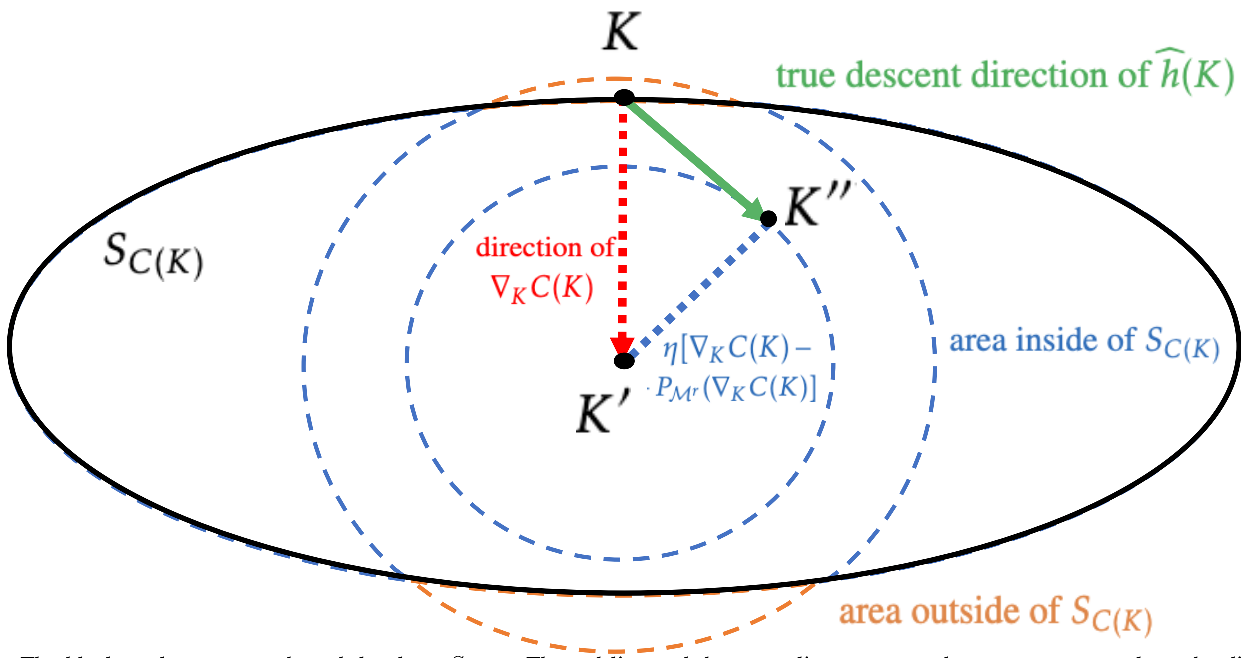


Fig. 1: The black oval represents the sub-level set $S_{C(\mathbf{K})}$. The red line and the green line represent the one-step move along the direction of $\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))$ and $\widehat{\mathbf{h}}(\mathbf{K}) = \sum_{i=1}^n \widehat{\mathbf{h}}_i(\mathbf{K})$, respectively. The blue line represents the difference caused by the gradient approximation and thus depends on κ. If the blue circle is so large that $C(\mathbf{K}'')$ moves out of $S_{C(\mathbf{K})}$ to the orange area, i.e., the κ is so small such that the approximation is too inaccurate, the system may take the risk of being unstable.

3) Guarantee of $C(\mathbf{K}'') \leq C(\mathbf{K})$: By combining the results in Appendix II-D2 and (7), we have,

$$
\begin{aligned}
& C(\mathbf{K}'') - C(\mathbf{K}) \\
=& C(\mathbf{K}'') - C(\mathbf{K}') + C(\mathbf{K}') - C(\mathbf{K}) \\
\leq& -(\eta - \frac{L}{2}\eta^2)\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\|^2 + \eta f_1^1(C(\mathbf{K}))\rho^{\kappa+1} + \eta^2 f_1^2(C(\mathbf{K}))\rho^{2(\kappa+1)} + \eta^3 f_1^3(C(\mathbf{K}))\rho^{3(\kappa+1)} \\
=& -\eta[-\eta^2 f_1^3(C(\mathbf{K}))\rho^{3(\kappa+1)} - \eta(f_1^2(C(\mathbf{K}))\rho^{2(\kappa+1)} + \frac{\mathbb{L}}{2}\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\|^2) + \|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\|^2 - f_1^1(C(\mathbf{K}))\rho^{\kappa+1}] \\
:=& -\eta(\varpi_2\eta^2 + \varpi_1\eta + \varpi_0),
\end{aligned}
$$

where

$$
\begin{aligned}
\varpi_2 &= -f_1^3(C(\mathbf{K}))\rho^{3(\kappa+1)}, \\
\varpi_1 &= -f_1^2(C(\mathbf{K}))\rho^{2(\kappa+1)} + \frac{\mathbb{L}}{2}\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\|, \\
\varpi_0 &= \|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\| - f_1^1(C(\mathbf{K}))\rho^{\kappa+1}.
\end{aligned}
$$

Note that $\varpi_2 < 0$. If we have $\kappa > \frac{1}{-\log\rho}\log\frac{2\sqrt{d}C(\mathbf{K})\mathcal{E}(C(\mathbf{K}))\mathbb{C}\sum_i L_i}{\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\|^2\sigma_1(\mathbf{Q})} - 1$, it is guaranteed that with all the η in the following range,

$$
0 < \eta \leq \frac{-\varpi_1 - \sqrt{\varpi_1^2 - 4\varpi_2\varpi_0}}{2\varpi_2}, \tag{12}
$$

the distributed gradient descent makes the objective function $C(\mathbf{K})$ decrease.

Remark 1. We have to emphasize that as the descent of the objective function, $\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\|$ may decrease, and the required low limit for κ may increase accordingly. Such a property is intuitive as shown in Fig. 1. When $\mathbf{K}$ gradually approaches the convergence point, the one-step improvement introduced by the exact gradient (the red dashed line in Fig. 1) may become smaller so that the tolerance to the approximation errors caused by the approximation error (the blue line) will become smaller as well. With a fixed communication range, the proposed method can only reach the neighborhood of the global optimal point. Otherwise, the move along the true descent direction (the green arrow in Fig. 1) will lead to the area outside of $S_{C(\mathbf{K})}$ and may cause the collapse of the stability. In this work, we operate under the assumption that the provided κ consistently meets the criteria outlined in Corollary **??** during the entire gradient process. The potential additional performance degradation occurring when this condition is unmet will be the subject of our future research endeavors.

E. Convergence and Degradation

Following the idea of decomposing $C(\mathbf{K}')-C(\mathbf{K})$ into $C(\mathbf{K}')-C(\mathbf{K}^h)$ and $C(\mathbf{K}^h)-C(\mathbf{K})$, we have that,

$$\begin{aligned}
&C(\mathbf{K}')-C(\mathbf{K})\\
=&\mathbb{E}_{\mathbf{x}_0\sim N(\mathbf{0},\boldsymbol{\Psi})}[\sum_{t\geq 0} A_{\mathbf{K},\mathbf{K}'}(\mathbf{x}'_t)]\\
=&2\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}(\mathbf{K}'-\mathbf{K})^T\mathbf{E}_{\mathbf{K}})+\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}(\mathbf{K}'-\mathbf{K})^T(\mathbf{R}+\mathbf{B}^T\mathbf{P}_{\mathbf{K}}\mathbf{B})(\mathbf{K}'-\mathbf{K}))\\
=&-2\eta\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}[\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}}C(\mathbf{K}))]^{\mathrm{T}}\mathbf{E}_{\mathbf{K}})+\eta^2\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}[\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}}C(\mathbf{K}))]^{\mathrm{T}}(\mathbf{R}+\mathbf{B}^T\mathbf{P}_{\mathbf{K}}\mathbf{B})[\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}}C(\mathbf{K}))])\\
=&\left[-2\eta\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}\nabla_{\mathbf{K}}{}^{\mathrm{T}}\mathbf{E}_{\mathbf{K}})+\eta^2\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}\nabla_{\mathbf{K}}{}^{\mathrm{T}}(\mathbf{R}+\mathbf{B}^T\mathbf{P}_{\mathbf{K}}\mathbf{B})\nabla_{\mathbf{K}})\right]\\
&+\left[2\eta\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}(\nabla_{\mathbf{K}}-\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}}))^T\mathbf{E}_{\mathbf{K}})+\eta^2\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}(\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}})-\nabla_{\mathbf{K}})^T(\mathbf{R}+\mathbf{B}^T\mathbf{P}_{\mathbf{K}}\mathbf{B})(\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}})-\nabla_{\mathbf{K}}))\right.\\
&\left.+\eta^2\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}(\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}})-\nabla_{\mathbf{K}})^T(\mathbf{R}+\mathbf{B}^T\mathbf{P}_{\mathbf{K}}\mathbf{B})\nabla_{\mathbf{K}})+\eta^2\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}\nabla_{\mathbf{K}}^{\mathrm{T}}(\mathbf{R}+\mathbf{B}^T\mathbf{P}_{\mathbf{K}}\mathbf{B})(\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}})-\nabla_{\mathbf{K}}))\right]\\
:=&\mathcal{B}_1+\mathcal{B}_2,
\end{aligned}$$

where $\mathcal{B}_1$ corresponds to the term $C(\mathbf{K}^h)-C(\mathbf{K})$ and $\mathcal{B}_2$ corresponds to the term $C(\mathbf{K}')-C(\mathbf{K}^h)$.

1) Detailed Discussion on $\mathcal{B}_1$: We have,

$$\begin{aligned}
&\mathcal{B}_1\\
=&-2\eta\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}\nabla_{\mathbf{K}}{}^{\mathrm{T}}\mathbf{E}_{\mathbf{K}})+\eta^2\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}\nabla_{\mathbf{K}}{}^{\mathrm{T}}(\mathbf{R}+\mathbf{B}^T\mathbf{P}_{\mathbf{K}}\mathbf{B})\nabla_{\mathbf{K}})\\
=&-2\eta\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}'}\boldsymbol{\Xi}_{\mathbf{K}}(\mathbf{E}_{\mathbf{K}})^{\mathrm{T}}\mathbf{E}_{\mathbf{K}})+\eta^2\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}}\boldsymbol{\Xi}_{\mathbf{K}'}\boldsymbol{\Xi}_{\mathbf{K}}(\mathbf{E}_{\mathbf{K}})^{\mathrm{T}}(\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B})\mathbf{E}_{\mathbf{K}})\\
\leq&-2\eta\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}}(\mathbf{E}_{\mathbf{K}})^{\mathrm{T}}\mathbf{E}_{\mathbf{K}}\boldsymbol{\Xi}_{\mathbf{K}})+2\eta\|\boldsymbol{\Xi}_{\mathbf{K}'}-\boldsymbol{\Xi}_{\mathbf{K}}\|\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}}(\mathbf{E}_{\mathbf{K}})^{\mathrm{T}}\mathbf{E}_{\mathbf{K}})+\eta^2\|\boldsymbol{\Xi}_{\mathbf{K}'}\|\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}}(\mathbf{E}_{\mathbf{K}})^{\mathrm{T}}\mathbf{E}_{\mathbf{K}}\boldsymbol{\Xi}_{\mathbf{K}})\\
\leq&-2\eta\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}}(\mathbf{E}_{\mathbf{K}})^{\mathrm{T}}\mathbf{E}_{\mathbf{K}}\boldsymbol{\Xi}_{\mathbf{K}})+2\eta\frac{\|\boldsymbol{\Xi}_{\mathbf{K}'}-\boldsymbol{\Xi}_{\mathbf{K}}\|}{\sigma_1(\boldsymbol{\Xi}_{\mathbf{K}})}\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}}(\mathbf{E}_{\mathbf{K}})^{\mathrm{T}}\mathbf{E}_{\mathbf{K}}\boldsymbol{\Xi}_{\mathbf{K}})+\eta^2\|\boldsymbol{\Xi}_{\mathbf{K}'}\|\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|\operatorname{Tr}(\boldsymbol{\Xi}_{\mathbf{K}}(\mathbf{E}_{\mathbf{K}})^{\mathrm{T}}\mathbf{E}_{\mathbf{K}}\boldsymbol{\Xi}_{\mathbf{K}})\\
=&-2\eta(1-\frac{\|\boldsymbol{\Xi}_{\mathbf{K}'}-\boldsymbol{\Xi}_{\mathbf{K}}\|}{\sigma_1(\boldsymbol{\Xi}_{\mathbf{K}})}-\frac{\eta}{2}\|\boldsymbol{\Xi}_{\mathbf{K}'}\|\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|)\operatorname{Tr}(\nabla C(\mathbf{K})^{\mathrm{T}}\nabla C(\mathbf{K}))\\
\leq&-2\eta\frac{\mu^2\sigma_1(\mathbf{R})}{\|\boldsymbol{\Xi}_{\mathbf{K}^*}\|}(1-\frac{\|\boldsymbol{\Xi}_{\mathbf{K}'}-\boldsymbol{\Xi}_{\mathbf{K}}\|}{\mu}-\frac{\eta}{2}\|\boldsymbol{\Xi}_{\mathbf{K}'}\|\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|)(C(\mathbf{K})-C(\mathbf{K}^*)),
\end{aligned}\tag{13}$$

where the last equation originates from the gradient domination property illustrated in Lemma 8. By lemma 11 and the condition of η in (15):

$$\frac{\|\boldsymbol{\Xi}_{\mathbf{K}'}-\boldsymbol{\Xi}_{\mathbf{K}}\|}{\mu}\leq 4\eta(\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})\mu})^2\|\mathbf{B}\|(\|\mathbf{A}-\mathbf{B}\mathbf{K}\|+1)\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}}C(\mathbf{K}))\|\leq 1/4,\tag{14}$$

where we demand that:

$$\eta\leq\frac{1}{16}(\frac{\sigma_1(\mathbf{Q})\mu}{C(\mathbf{K})})^2\frac{1}{\|\mathbf{B}\|\|\mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}}C(\mathbf{K}))\|(1+\|\mathbf{A}-\mathbf{B}\mathbf{K}\|)}.\tag{15}$$

Combining (14) and Lemma 12.

$$\|\boldsymbol{\Xi}_{\mathbf{K}'}\|\leq\|\boldsymbol{\Xi}_{\mathbf{K}'}-\boldsymbol{\Xi}_{\mathbf{K}}\|+\|\boldsymbol{\Xi}_{\mathbf{K}}\|\leq\frac{\mu}{4}+\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})}\leq\frac{\|\boldsymbol{\Xi}_{\mathbf{K}'}\|}{4}+\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})},$$

so that $\|\boldsymbol{\Xi}_{\mathbf{K}'}\|\leq\frac{4C(\mathbf{K})}{3\sigma_1(\mathbf{Q})}$. Along with the condition on η in (1).

$$1-\frac{\|\boldsymbol{\Xi}_{\mathbf{K}'}-\boldsymbol{\Xi}_{\mathbf{K}}\|}{\mu}-\frac{\eta}{2}\|\boldsymbol{\Xi}_{\mathbf{K}'}\|\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|\geq 1-1/4-\frac{\eta}{2}\frac{4C(\mathbf{K})}{3\sigma_1(\mathbf{Q})}\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|\geq 1/2,$$

where we demand that:

$$\eta\leq\frac{\sigma_1(\mathbf{Q})}{32C(\mathbf{K})\|\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B}\|},\tag{16}$$

so that

$$\mathcal{B}_1\leq-\eta\frac{\mu^2\sigma_1(\mathbf{R})}{\|\boldsymbol{\Xi}_{\mathbf{K}^*}\|}(C(\mathbf{K})-C(\mathbf{K}^*)).\tag{17}$$

Remark 2. Combining (3), (6), (12), (15) and (16), we obtain the full condition for η as shown in (1).

Remark 3. We emphasize that every term in the conditions of step size can be bounded uniformly in the gradient process. To see that, we observe that $C(\mathbf{K}^*) \leq C(\mathbf{K}) \leq C(\mathbf{K}(0))$, $\Upsilon(C(\mathbf{K}^*)) \leq \Upsilon(C(\mathbf{K})) \leq \Upsilon(C(\mathbf{K}(0)))$, $\mathcal{E}(C(\mathbf{K}^*)) \leq \mathcal{E}(C(\mathbf{K})) \leq \mathcal{E}(C(\mathbf{K}(0)))$. $\|\mathbf{K}\|$(Lemma 10),$\|\nabla_{\mathbf{K}}\|$ (Lemma 10), $\|\mathbf{P}_{\mathbf{K}}\|$ (Lemma 12) can be all bounded by the polynomial regarding $C(\mathbf{K})$. So that

$$\eta = \textbf{poly}(\frac{1}{C(\mathbf{K}(0))}, \frac{1}{\|\mathbf{A}\|}, \frac{1}{\|\mathbf{B}\|}, \frac{1}{\|\mathbf{Q}\|}, \sigma_1(\mathbf{Q}), \sigma_1(\mathbf{R}), \mu, \frac{1}{C\sum_i L_i}, \frac{1}{\rho^{\kappa+1}}, \sigma_n(\mathbf{Q}), \sigma_1(\mathbf{\Psi}), \frac{1}{\sqrt{d}}).$$

2) Detailed Discussion on $\mathcal{B}_2$: First, we present the SED property for $\nabla_{\mathbf{K}} C(\mathbf{K})$.

Lemma 5. If Lemma IV.16 or Lemma IV.17 holds, $\nabla_{\mathbf{K}} C(\mathbf{K})$ is $(C_{\nabla_{\mathbf{K}}}, \rho)$ spatially exponential decaying (SED).

The definition of SED is illustrated in Definition 1. The value of $C_{\nabla_{\mathbf{K}}}$ and the detailed proof are deferred to Section IV-C. Then we have

$$\begin{aligned}
&\left\|\nabla_{\mathbf{K}} C(\mathbf{K}) - \mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\right\| \leq \left\|\nabla_{\mathbf{K}} C(\mathbf{K}) - \mathcal{P}_{\mathcal{M}^r}(\nabla_{\mathbf{K}} C(\mathbf{K}))\right\|_F \leq \sum_{i=1}^{n} \sum_{j \in \mathcal{N}_i^{-r}} \left\|[\nabla_{\mathbf{K}} C(\mathbf{K})]_{ij}\right\|_F \\
&\leq \sum_{i=1}^{n} \sum_{j \in \mathcal{N}_i^{-r}} C_{\nabla_{\mathbf{K}}} \rho^{\mathrm{dist}(i,j)} \leq \sum_{i=1}^{n} |\mathcal{N}_i^{-r}| C_{\nabla_{\mathbf{K}}} \rho^r \leq n \max_i (|\mathcal{N}_i^{-r}|) C_{\nabla_{\mathbf{K}}} \rho^r.
\end{aligned}$$

We observe that this term decays to 0 exponentially as κ and r increase. With (10), (8) and Lemma 12, we have,

$$\begin{aligned}
\mathcal{B}_2 \leq & \frac{2\eta C(\mathbf{K})}{\sigma_1(\mathbf{Q})} \mathcal{E}(C(\mathbf{K})) n \max_i (|\mathcal{N}_i^{-r}|) \rho^r + \eta^2 \frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})} \left[n \max_i (|\mathcal{N}_i^{-r}|) \rho^r\right]^2 \left(\|\mathbf{R}\| + \|\mathbf{B}\|^2 \frac{C(\mathbf{K})}{\mu}\right) \\
& + 2\eta^2 \frac{C(\mathbf{K})^2}{(\sigma_1(\mathbf{Q}))^2} \left[n \max_i (|\mathcal{N}_i^{-r}|) \rho^r\right] \left(\|\mathbf{R}\| + \|\mathbf{B}\|^2 \frac{C(\mathbf{K})}{\mu}\right) \\
:= & M_2 F_2(C(\mathbf{K})) \rho^r.
\end{aligned} \tag{18}$$

Combining the results in (17), (18) and (11), we have,

$$C(\mathbf{K}'') - C(\mathbf{K}) \leq -\eta \frac{\mu^2 \sigma_1(\mathbf{R})}{\|\mathbf{\Xi}_{\mathbf{K}^*}\|} (C(\mathbf{K}) - C(\mathbf{K}^*)) + M_1 F_1(C(\mathbf{K})) \rho^{\kappa+1} + M_2 F_2(C(\mathbf{K}))) \rho^r,$$

where we use $M_1 F_1(\cdot) \rho^{\kappa+1}$ to denote $\eta f_1^1(\cdot) \rho^{\kappa+1} + \eta^2 f_1^2(\cdot) \rho^{2(\kappa+1)} + \eta^3 f_1^3(\cdot) \rho^{3(\kappa+1)}$.

Remark 4. Both $F_1(C(\mathbf{K}))$ and $F_2(C(\mathbf{K}))$ monotonically increases with the increase of $C(\mathbf{K})$, so that $F_1(C(\mathbf{K}^*) \leq F_1(C(\mathbf{K}(T))) \leq F_1(C(\mathbf{K}(0)))$, $F_2(C(\mathbf{K}^*) \leq F_2(\mathbf{K}(T)) \leq F_2(C(\mathbf{K}(0)))$, for any t from 1 to T.

Thus,

$$\begin{aligned}
& C(\mathbf{K}(T)) - C(\mathbf{K}^*) \\
& \leq \left(1 - \eta \frac{\mu^2 \sigma_1(\mathbf{R})}{\|\mathbf{\Xi}_{\mathbf{K}^*}\|}\right) \left[C(\mathbf{K}(t-1)) - C(\mathbf{K}^*)\right] + M_1 F_1(C(\mathbf{K}(t-1))) \rho^{\kappa+1} + M_2 F_2(C(\mathbf{K}(t-1))) \rho^r \\
& \leq \left(1 - \eta \frac{\mu^2 \sigma_1(\mathbf{R})}{\|\mathbf{\Xi}_{\mathbf{K}^*}\|}\right) \left[C(\mathbf{K}(t-1)) - C(\mathbf{K}^*)\right] + M_1 F_1(C(\mathbf{K}(0))) \rho^{\kappa+1} + M_2 F_2(C(\mathbf{K}(0))) \rho^r.
\end{aligned}$$

Inductively, we have,

$$\begin{aligned}
& C(\mathbf{K}(T)) - C(\mathbf{K}^*) \\
= & [C(\mathbf{K}(T)) - C(\mathbf{K}(T-1)] + [C(\mathbf{K}(T-1) - C(\mathbf{K}^*)] \\
\leq & (1 - \eta \frac{\mu^2 \sigma_1(\mathbf{R})}{\|\mathbf{\Xi}_{\mathbf{K}^*}\|}) \left[C(\mathbf{K}(T-1)) - C(\mathbf{K}^*)\right] + M_1 F_1(C(\mathbf{K}(t-1))) \rho^{\kappa+1} + M_2 F_2(C(\mathbf{K}(t-1))) \rho^r \\
\leq & (1 - \eta \frac{\mu^2 \sigma_1(\mathbf{R})}{\|\mathbf{\Xi}_{\mathbf{K}^*}\|}) \left[C(\mathbf{K}(T-1)) - C(\mathbf{K}^*)\right] + M_1 F_1(C(\mathbf{K}(0))) \rho^{\kappa+1} + M_2 F_2(C(\mathbf{K}(0))) \rho^r \\
\leq & (1 - \eta \frac{\mu^2 \sigma_1(\mathbf{R})}{\|\mathbf{\Xi}_{\mathbf{K}^*}\|})^T \left[C(\mathbf{K}(0)) - C(\mathbf{K}^*)\right] + \prod_{\tau=0}^{T-1} (1 - \eta \frac{\mu^2 \sigma_1(\mathbf{R})}{\|\mathbf{\Xi}_{\mathbf{K}^*}\|})^\tau [M_1 F_1(C(\mathbf{K}(0))) \rho^{\kappa+1} + M_2 F_2(C(\mathbf{K}(0))) \rho^r] \\
\leq & (1 - \eta \frac{\mu^2 \sigma_1(\mathbf{R})}{\|\mathbf{\Xi}_{\mathbf{K}^*}\|})^T (C(\mathbf{K}(0)) - C(\mathbf{K}^*)) + \frac{\|\mathbf{\Xi}_{\mathbf{K}^*}\|}{\eta \mu^2 \sigma_1(\mathbf{R})} [M_1 F_1(C(\mathbf{K}(0))) \rho^{\kappa+1} + M_2 F_2(C(\mathbf{K}(0))) \rho^r],
\end{aligned}$$

Provided

$$T \geq \frac{\|\mathbf{\Xi}_{\mathbf{K}^*}\|}{\eta\mu^2\sigma_1(\mathbf{R})} \log \frac{C(\mathbf{K}(0)) - C(\mathbf{K}^*)}{\epsilon},$$

then we have

$$C(\mathbf{K}(T)) - C(\mathbf{K}^*) \leq \epsilon + \frac{\|\mathbf{\Xi}_{\mathbf{K}^*}\|}{\eta\mu^2\sigma_1(\mathbf{R})}\big[M_1F_1(C(\mathbf{K}(0)))\rho^{\kappa+1} + M_2F_2(C(\mathbf{K}(0)))\rho^{r}\big].$$

The proof of Theorem IV.3 is done.

III. Revisit LQ Setting from the Perspective of Single Agent

Lemma 6. [5] For any $\mathbf{K}$ such that $\rho(\mathbf{A}-\mathbf{BK}) < 1$. Let $\mathbf{P_K}$ be the unique positive definite solution to the Bellman equation

$$\mathbf{P_K} = (\mathbf{Q} + \mathbf{K}^{\mathrm{T}}\mathbf{RK}) + (\mathbf{A}-\mathbf{BK})^{\mathrm{T}}\mathbf{P_K}(\mathbf{A}-\mathbf{BK}).$$

In the setting of LQR, we have

$$\begin{aligned}
V_{\mathbf{K}}(\mathbf{x}) =& \mathbf{x}^{\mathrm{T}}\mathbf{P_K}\mathbf{x} - \operatorname{Tr}(\mathbf{P_K}\mathbf{\Xi_K}) \\
C(\mathbf{K}) =& \operatorname{Tr}[(\mathbf{Q}+\mathbf{K}^{\mathrm{T}}\mathbf{RK})\mathbf{\Xi_K}] + \sigma_0^2\operatorname{Tr}(\mathbf{R}) = \operatorname{Tr}(\mathbf{P_K}\mathbf{\Psi}) + \sigma_0^2\operatorname{Tr}(\mathbf{R}) \\
Q_{\mathbf{K}}(\mathbf{x},\mathbf{u}) =& \mathbf{x}^{\mathrm{T}}(\mathbf{Q}+\mathbf{A}^{\mathrm{T}}\mathbf{P_K}\mathbf{A})\mathbf{x} + \mathbf{x}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{P_K}\mathbf{B})\mathbf{u} + \mathbf{u}^{\mathrm{T}}(\mathbf{B}^{\mathrm{T}}\mathbf{P_K}\mathbf{A})\mathbf{x} \\
&+ \mathbf{u}^{\mathrm{T}}(\mathbf{R}+\mathbf{B}^{\mathrm{T}}\mathbf{P_K}\mathbf{B})u - \operatorname{Tr}(\mathbf{P_K}\mathbf{\Xi_K}) - {\sigma_0}^2\operatorname{Tr}(\mathbf{R}+\mathbf{P_K}\mathbf{BB}^{\mathrm{T}}).
\end{aligned}$$

Proof. Note that under $\pi_{\mathbf{K}}$, we can write $\mathbf{u}_t$ as $-\mathbf{K}\mathbf{x}_t + \sigma_0 \cdot \eta_t$, where $\eta \sim N(\mathbf{0}, \boldsymbol{I}_{d_u})$. For all $t \geq 0$, we have

$$\mathbb{E}(c(t)|\mathbf{x}_t) = \mathbf{x}_t^{\mathrm{T}}\mathbf{Q}\mathbf{x}_t + \mathbb{E}_{\eta_t\sim N(\mathbf{0},\boldsymbol{I}_{d_u})}[(-\mathbf{K}\mathbf{x}_t+\sigma_0\eta_t)^{\mathrm{T}}R(-\mathbf{K}\mathbf{x}_t+\sigma_0\eta_t)] = \mathbf{x}_t^{\mathrm{T}}(\mathbf{Q}+\mathbf{K}^{\mathrm{T}}\mathbf{RK})\mathbf{x}_t + \sigma_0^2\operatorname{Tr}(\mathbf{R}).$$

Thus, according to the definition of $C(\mathbf{K})$, we have

$$\begin{aligned}
C(\mathbf{K}) &= \lim_{T\to\infty}\mathbb{E}\{\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[c(\mathbf{x}_t,\mathbf{u}_t)|x_t]\} \\
&= \lim_{T\to\infty}\mathbb{E}\{\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\mathbf{x}_t^{\mathrm{T}}(\mathbf{Q}+\mathbf{K}^{\mathrm{T}}\mathbf{RK})\mathbf{x}_t + \sigma_0^2\operatorname{Tr}(\mathbf{R})]\} \\
&= \mathbb{E}_{\mathbf{x}\sim\nu_{\mathbf{K}}}[\mathbf{x}^{\mathrm{T}}(\mathbf{Q}+\mathbf{K}^{\mathrm{T}}\mathbf{RK})\mathbf{x} + \sigma_0^2\operatorname{Tr}(\mathbf{R})] \\
&= \operatorname{Tr}[(\mathbf{Q}+\mathbf{K}^{\mathrm{T}}\mathbf{RK})\mathbf{\Xi_K}] + \sigma_0^2\operatorname{Tr}(\mathbf{R}).
\end{aligned}$$

We have $\mathbf{\Xi_K} = \mathcal{T}_{\mathbf{K}}(\mathbf{\Psi})$ and $\mathbf{P_K} = \mathcal{T}_{\mathbf{K}}^{\mathrm{T}}(\mathbf{Q}+\mathbf{K}^{\mathrm{T}}\mathbf{RK})$. Thus,

$$\operatorname{Tr}[(\mathbf{Q}+\mathbf{K}^{\mathrm{T}}\mathbf{RK})\mathbf{\Xi_K}] = \operatorname{Tr}[(\mathbf{Q}+\mathbf{K}^{\mathrm{T}}\mathbf{RK})\mathcal{T}_{\mathbf{K}}(\mathbf{\Psi})] = \operatorname{Tr}[\mathcal{T}_{\mathbf{K}}^{\mathrm{T}}(\mathbf{Q}+\mathbf{K}^{\mathrm{T}}\mathbf{RK})\mathbf{\Psi}] = \operatorname{Tr}(\mathbf{P_K}\mathbf{\Psi}).$$

In the setting of LQR, the state-value function $V_{\mathbf{K}}$ is given by

$$V_{\mathbf{K}}(\mathbf{x}) = \sum_{t=0}^{\infty}\{\mathbb{E}[\mathbf{x}_t^{\mathrm{T}}(\mathbf{Q}+\mathbf{K}^{\mathrm{T}}\mathbf{RK})\mathbf{x}_t] + \sigma_0^2\operatorname{Tr}(\mathbf{R}) - C(\mathbf{K})\}.$$

Combining the linear dynamics, we see that $V_{\mathbf{K}}$ is a quadratic function. We denote it by $V_{\mathbf{K}}(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{P_K}\mathbf{x} + \alpha_{\mathbf{K}}$, where both $\mathbf{P_K}$ and $\alpha_{\mathbf{K}}$ depend on $\mathbf{K}$. By Bellman equation:

$$V_{\mathbf{K}}(\mathbf{x}) = \mathbb{E}_{u\sim\pi_{\mathbf{K}}}[c(\mathbf{x},\mathbf{u})] - C(\mathbf{K}) + \mathbb{E}[V_{\mathbf{K}}(\mathbf{x}')|\mathbf{x}],$$

thus, for any $\mathbf{x} \in \mathbb{R}^d$, we have

$$\mathbf{x}^{\mathrm{T}}\mathbf{P_K}\mathbf{x} = \mathbf{x}^{\mathrm{T}}(\mathbf{Q}+\mathbf{K}^{\mathrm{T}}\mathbf{RK})\mathbf{x} + \mathbf{x}^{\mathrm{T}}(\mathbf{A}-\mathbf{BK})^{\mathrm{T}}\mathbf{P_K}(\mathbf{A}-\mathbf{BK})\mathbf{x}.$$

So that $\mathbf{P_K}$ is the unique positive definite solution to the Bellman equation. Meanwhile, since $\mathbb{E}_{\mathbf{x}\sim\nu_{\mathbf{K}}}V_{\mathbf{K}}(\mathbf{x}) = 0$, we have $\alpha_{\mathbf{K}} = -\operatorname{Tr}(\mathbf{P_K}\mathbf{\Xi_K})$.

Furthermore, for any state-action pair $(\mathbf{x}, \mathbf{u})$, we have,

$$\begin{aligned}
Q_{\mathbf{K}}(\mathbf{x}, \mathbf{u}) =& c(\mathbf{x}, \mathbf{u}) - C(\mathbf{K}) + \mathbb{E}[V_{\mathbf{K}}(\mathbf{x}')|\mathbf{x}, \mathbf{u}] \\
=& c(\mathbf{x}, \mathbf{u}) - C(\mathbf{K}) + \mathbb{E}[(\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{w})^{\mathrm{T}} \mathbf{P}_{\mathbf{K}} (\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{w})|\mathbf{x}, \mathbf{u}] - \mathrm{Tr}(\mathbf{P}_{\mathbf{K}} \mathbf{\Xi}_{\mathbf{K}}) \\
=& c(\mathbf{x}, \mathbf{u}) - C(\mathbf{K}) + (\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u})^{\mathrm{T}} \mathbf{P}_{\mathbf{K}} (\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}) + \mathrm{Tr}(\mathbf{P}_{\mathbf{K}} \mathbf{\Phi}) - \mathrm{Tr}(\mathbf{P}_{\mathbf{K}} \mathbf{\Xi}_{\mathbf{K}}) \\
=& \mathbf{x}^{\mathrm{T}} \mathbf{Q} \mathbf{x} + \mathbf{u}^{\mathrm{T}} \mathbf{R} \mathbf{u} + (\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u})^{\mathrm{T}} \mathbf{P}_{\mathbf{K}} (\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}) - \sigma_0^2 \mathrm{Tr}(\mathbf{R} + \mathbf{P}_{\mathbf{K}} \mathbf{B}\mathbf{B}^{\mathrm{T}}) - \mathrm{Tr}(\mathbf{P}_{\mathbf{K}} \mathbf{\Xi}_{\mathbf{K}}).
\end{aligned}$$

□

Lemma 7. [6] Let $\mathbf{K}$ and $\mathbf{K}'$ be two controllers that $\rho(\mathbf{A} - \mathbf{B}\mathbf{K}) < 1$ and $\rho(\mathbf{A} - \mathbf{B}\mathbf{K}') < 1$. For any $\mathbf{x} \in \mathbb{R}^d$, let $\{\mathbf{x}'_t\}_{t \geq 0}$ be the sequence of states satisfying $\mathbf{x}'_0 = \mathbf{x}$ and $\mathbf{x}'_{t+1} = (\mathbf{A} - \mathbf{B}\mathbf{K}')\mathbf{x}'_t$ for all $t \geq 0$. Then it holds that

$$\mathbf{x}^T \mathbf{P}_{\mathbf{K}'} \mathbf{x} - \mathbf{x}^T \mathbf{P}_{\mathbf{K}} \mathbf{x} = \sum_{t \geq 0} A_{\mathbf{K}, \mathbf{K}'}(\mathbf{x}'_t),$$

where $A_{\mathbf{K},\mathbf{K}'}(x) = 2\mathbf{x}^{\mathrm{T}}(\mathbf{K}' - \mathbf{K})^{\mathrm{T}} \mathbf{E}_{\mathbf{K}} \mathbf{x} + \mathbf{x}^{\mathrm{T}} (\mathbf{K}' - \mathbf{K})^{\mathrm{T}} (\mathbf{R} + \mathbf{B}^{\mathrm{T}} \mathbf{P}_{\mathbf{K}} \mathbf{B})(\mathbf{K}' - \mathbf{K})\mathbf{x}$ and $\mathbf{E}_{\mathbf{K}} = (\mathbf{R} + \mathbf{B}^{\mathrm{T}} \mathbf{P}_{\mathbf{K}} \mathbf{B})\mathbf{K} - \mathbf{B}^{\mathrm{T}} \mathbf{P}_{\mathbf{K}} \mathbf{A}$.

Lemma 8. [6] (Gradient Domination) Let $\mathbf{K}^*$ be an optimal policy. Suppose $\mathbf{K}$ has a finite cost. It holds that,

$$C(\mathbf{K}) - C(\mathbf{K}^*) \leq \frac{\|\mathbf{\Xi}_{\mathbf{K}^*}\|}{\sigma_1(\mathbf{\Xi}_{\mathbf{K}})^2 \sigma_1(\mathbf{R})} \|\nabla_{\mathbf{K}} C(\mathbf{K})\|_F^2.$$

Using the fact that $\mathbf{\Xi}_{\mathbf{K}} \succeq \mathbf{\Psi}$, the following corollary shows that $C(\mathbf{K})$ is gradient dominated.

Corollary 1. [6] Suppose $\mathbb{E}_{\mathbf{x}_0 \sim N(\mathbf{0}, \mathbf{\Psi})} \mathbf{x}_0 \mathbf{x}_0^{\mathrm{T}}$ is full rank. Then $C(\mathbf{K})$ is gradient dominated, i.e.

$$C(\mathbf{K}) - C(\mathbf{K}^*) \leq \lambda \left\langle \nabla_{\mathbf{K}} C(\mathbf{K}), \nabla_{\mathbf{K}} C(\mathbf{K}) \right\rangle,$$

where $\lambda = \frac{\|\mathbf{\Xi}_{\mathbf{K}^*}\|}{\sigma_1(\mathbf{\Xi}_0)^2 \sigma_1(\mathbf{R})}$ is a problem dependent constant and $\langle \cdot, \cdot \rangle$ denotes the trace inner product.

Lemma 9. [6] Let $\mathbf{K}^*$ be an optimal policy. Suppose $\mathbf{K}$ has a finite cost. It holds that:

$$C(\mathbf{K}) - C(\mathbf{K}^*) \geq \frac{\mu}{\|\mathbf{R} + \mathbf{B}^{\mathrm{T}} \mathbf{P}_{\mathbf{K}} \mathbf{B}\|} \mathrm{Tr}(\mathbf{E}_{\mathbf{K}}^{\mathrm{T}} \mathbf{E}_{\mathbf{K}}).$$

Lemma 10. [6] It holds that

$$\begin{aligned}
\|\mathbf{K}\| &\leq \frac{1}{\sigma_1(\mathbf{R})} \left(\sqrt{\frac{\|\mathbf{R} + \mathbf{B}^{\mathrm{T}} \mathbf{P}_{\mathbf{K}} \mathbf{B}\| (C(\mathbf{K}) - C(\mathbf{K}^*))}{\mu}} + \|\mathbf{B}^{\mathrm{T}} \mathbf{P}_{\mathbf{K}} \mathbf{A}\| \right), \\
\|\nabla C(\mathbf{K})\| &\leq \frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})} \sqrt{\frac{\|\mathbf{R} + \mathbf{B}^{\mathrm{T}} \mathbf{P}_{\mathbf{K}} \mathbf{B}\| (C(\mathbf{K}) - C(\mathbf{K}^*))}{\mu}}.
\end{aligned}$$

Lemma 11. [6] ($\mathbf{\Xi}_{\mathbf{K}}$ perturbation) Suppose $\mathbf{K}'$ is such that:

$$\|\mathbf{K}' - \mathbf{K}\| \leq \frac{\sigma_1(\mathbf{Q})\mu}{4C(\mathbf{K})\|\mathbf{B}\|(\|\mathbf{A} - \mathbf{B}\mathbf{K}\| + 1)}.$$

It holds that $\rho(\mathbf{A} - \mathbf{B}\mathbf{K}') < 1$ and,

$$\|\mathbf{\Xi}_{\mathbf{K}'} - \mathbf{\Xi}_{\mathbf{K}}\| \leq 4 (\frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})})^2 \frac{\|\mathbf{B}\|(\|\mathbf{A} - \mathbf{B}\mathbf{K}\| + 1)}{\mu} \|\mathbf{K}' - \mathbf{K}\|.$$

Lemma 12. [6] ($\mathbf{P}_{\mathbf{K}}$ and $\mathbf{\Xi}_{\mathbf{K}}$ norm bound) For a stabilizing K, it holds that:

$$\|\mathbf{P}_{\mathbf{K}}\| \leq \frac{C(\mathbf{K})}{\mu}, \quad \|\mathbf{\Xi}_{\mathbf{K}}\| \leq \frac{C(\mathbf{K})}{\sigma_1(\mathbf{Q})}. \tag{19}$$

Lemma 13. [3] On the sub-level set $S_{C(\mathbf{K}(0))}$, the gradient $\nabla C(\mathbf{K})$ is L-Lipschitz continuous and we have:

$$\begin{aligned}
& \|\nabla^2 C(\mathbf{K})\| && (20) \\
\leq & \left(2\sigma_n(\mathbf{R}) + \frac{2\|\mathbf{B}\|^2 C(\mathbf{K}(0))}{\sigma_1(\mathbf{\Xi}_{\mathbf{K}})} + 4\sqrt{2}\zeta \|\mathbf{B}\| \frac{C(\mathbf{K}(0))}{\sigma_1(\mathbf{\Xi}_{\mathbf{K}})} \right) \frac{C(\mathbf{K}(0))}{\sigma_n(\mathbf{Q})} && (21) \\
\leq & \left(2\sigma_n(\mathbf{R}) + \frac{2\|\mathbf{B}\|^2 C(\mathbf{K}(0))}{\sigma_1(\mathbf{\Psi})} + 4\sqrt{2}\zeta \|\mathbf{B}\| \frac{C(\mathbf{K}(0))}{\mu} \right) \frac{C(\mathbf{K}(0))}{\sigma_n(\mathbf{Q})} && (22) \\
:= & \mathbb{L}, && (23)
\end{aligned}$$

where we have $\mathbf{\Xi}_{\mathbf{K}} \succeq \mathbf{\Psi}$ and ζ is the constant that only determined by problem data ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{Q}$, $\mathbf{R}$, μ, $\mathbf{K}(0)$)

A. Proof of Lemma IV.2

Proof. The proof is similar to the tabular case in [7], but we generalize it from the tabular form to the continuous space. First, we prove part (a) that the truncated local Q function $\mathbb{Q}^i_{\mathbf{K},\kappa}(\mathbf{x}_{\mathcal{N}_i^\kappa},\mathbf{u}_{\mathcal{N}_i^\kappa})$ is a good approximation of the exact local Q function $Q^i_{\mathbf{K}}(\mathbf{x},\mathbf{u})$, which is shown in (24).

$$
\begin{aligned}
&\left|\mathbb{Q}^i_{\mathbf{K},\kappa}(\mathbf{x}_{\mathcal{N}_i^\kappa},\mathbf{u}_{\mathcal{N}_i^\kappa}) - Q^i_{\mathbf{K}}(\mathbf{x},\mathbf{u})\right| &(24)\\
&=\left|\int_{\mathbf{x}'_{\mathcal{N}^\kappa_{-i}}}\int_{\mathbf{u}'_{\mathcal{N}^\kappa_{-i}}} \zeta_i(\mathbf{x}'_{\mathcal{N}^\kappa_{-i}},\mathbf{u}'_{\mathcal{N}^\kappa_{-i}};\mathbf{x}_{\mathcal{N}_i^\kappa},\mathbf{u}_{\mathcal{N}_i^\kappa})Q^i_{\mathbf{K}}(\mathbf{x}_{\mathcal{N}_i^\kappa},\mathbf{x}'_{\mathcal{N}^\kappa_{-i}},\mathbf{u}_{\mathcal{N}_i^\kappa},\mathbf{u}'_{\mathcal{N}^\kappa_{-i}})d\mathbf{x}'_{\mathcal{N}^\kappa_{-i}}d\mathbf{u}'_{\mathcal{N}^\kappa_{-i}} - Q^i_{\mathbf{K}}(\mathbf{x}_{\mathcal{N}_i^\kappa},\mathbf{x}_{\mathcal{N}^\kappa_{-i}},\mathbf{u}_{\mathcal{N}_i^\kappa},\mathbf{u}_{\mathcal{N}^\kappa_{-i}})\right| &(25)\\
&\le \int_{\mathbf{x}'_{\mathcal{N}^\kappa_{-i}}}\int_{\mathbf{u}'_{\mathcal{N}^\kappa_{-i}}} \zeta_i(\mathbf{x}'_{\mathcal{N}^\kappa_{-i}},\mathbf{u}'_{\mathcal{N}^\kappa_{-i}};\mathbf{x}_{\mathcal{N}_i^\kappa},\mathbf{u}_{\mathcal{N}_i^\kappa})d\mathbf{x}'_{\mathcal{N}^\kappa_{-i}}d\mathbf{u}'_{\mathcal{N}^\kappa_{-i}}\left|Q^i_{\mathbf{K}}(\mathbf{x}_{\mathcal{N}_i^\kappa},\mathbf{x}'_{\mathcal{N}^\kappa_{-i}},\mathbf{u}_{\mathcal{N}_i^\kappa},\mathbf{u}'_{\mathcal{N}^\kappa_{-i}}) - Q^i_{\mathbf{K}}(\mathbf{x}_{\mathcal{N}_i^\kappa},\mathbf{x}_{\mathcal{N}^\kappa_{-i}},\mathbf{u}_{\mathcal{N}_i^\kappa},\mathbf{u}_{\mathcal{N}^\kappa_{-i}})\right| &(26)\\
&\le C\rho^{\kappa+1}. &(27)
\end{aligned}
$$

Next, we show part (b). Recall the policy gradient theorem (Lemma III.2)

$$\nabla_{\mathbf{K}_i}C(\mathbf{K}) = \mathbb{E}_{\mathbf{x}\sim\nu_{\mathbf{K}},\mathbf{u}\sim\pi_{\mathbf{K}}(\cdot|x)}\left[Q_{\mathbf{K}}(\mathbf{x},\mathbf{u})\nabla_{\mathbf{K}_i}\log\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})\right],$$

where we use the localized policy $\pi_{\mathbf{K}}(\mathbf{u}|\mathbf{x}) = \prod_{i=1}^N \pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})$,

With the above equation, we can compute $\widehat{\mathbf{h}}_i(\mathbf{K}) - \nabla_{\mathbf{K}_i}C(\mathbf{K})$,

$$
\begin{aligned}
&\widehat{\mathbf{h}}_i(\mathbf{K}) - \nabla_{\mathbf{K}_i}C(\mathbf{K})\\
=&\mathbb{E}_{\mathbf{x}\sim\nu_{\mathbf{K}},\mathbf{u}\sim\pi_{\mathbf{K}}(\cdot|\mathbf{x})}\Big[\frac{1}{n}\sum_{j\in\mathcal{N}_i^\kappa}\mathbb{Q}^j_{\mathbf{K},\kappa}(\mathbf{x}_{\mathcal{N}_j^\kappa},\mathbf{u}_{\mathcal{N}_j^\kappa}) - Q_{\mathbf{K}}(\mathbf{x},\mathbf{u})\Big]\nabla_{\mathbf{K}_i}\log\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})\\
=&\mathbb{E}_{\mathbf{x}\sim\nu_{\mathbf{K}},\mathbf{u}\sim\pi_{\mathbf{K}}(\cdot|\mathbf{x})}\Big[\frac{1}{n}\sum_{j=1}^{n}\mathbb{Q}^j_{\mathbf{K},\kappa}(\mathbf{x}_{\mathcal{N}_j^\kappa},\mathbf{u}_{\mathcal{N}_j^\kappa}) - \frac{1}{n}\sum_{j=1}^{n}Q^i_{\mathbf{K}}(\mathbf{x},\mathbf{u})\Big]\nabla_{\mathbf{K}_i}\log\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})\\
&-\mathbb{E}_{\mathbf{x}\sim\nu_{\mathbf{K}},\mathbf{u}\sim\pi_{\mathbf{K}}(\cdot|\mathbf{x})}\sum_{j\in\mathcal{N}^\kappa_{-i}}\frac{1}{n}\mathbb{Q}^j_{\mathbf{K},\kappa}(\mathbf{x}_{\mathcal{N}_j^\kappa},\mathbf{u}_{\mathcal{N}_j^\kappa})\nabla_{\mathbf{K}_i}\log\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})\\
:=&E_1 - E_2.
\end{aligned}
$$

We claim that $E_2 = 0$. To see this, we use $d^{\mathbf{K}}(\mathbf{x})$ to denote the stationary distribution over $\mathbf{x}$, where $d^{\mathbf{K}}(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}|\mathbf{\Xi}_{\mathbf{K}}|}\exp\left(-\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{\Xi}_{\mathbf{K}}^{-1}x\right)$. Then for any $j\in\mathcal{N}^\kappa_{-i}$, we have:

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{x}\sim\nu_{\mathbf{K}},\mathbf{u}\sim\pi_{\mathbf{K}}(\cdot|\mathbf{x})}\mathbb{Q}^j_{\mathbf{K},\kappa}(\mathbf{x}_{\mathcal{N}_j^\kappa},\mathbf{u}_{\mathcal{N}_j^\kappa})\nabla_{\mathbf{K}_i}\log\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})\\
&=\int_{\mathbf{x},\mathbf{u}} d^{\mathbf{K}}(\mathbf{x})\prod_{l=1}^{n}\pi_{K_l}(u_l|x_{N_l^r})\frac{\nabla_{\mathbf{K}_i}\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})}{\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})}\mathbb{Q}^K_j(\mathbf{x}_{\mathcal{N}_j^\kappa},\mathbf{u}_{\mathcal{N}_j^\kappa})d\mathbf{x}d\mathbf{u}\\
&=\int_{\mathbf{x},\mathbf{u}_1,\ldots,\mathbf{u}_{i-1},\mathbf{u}_{i+1},\ldots,\mathbf{u}_n} d^{\mathbf{K}}(\mathbf{x})\prod_{l\neq i}\pi_{\mathbf{K}_l}(\mathbf{u}_l|\mathbf{x}_{N_l^r})\mathbb{Q}^{\mathbf{K}}_j(\mathbf{x}_{\mathcal{N}_j^\kappa},\mathbf{u}_{\mathcal{N}_j^\kappa})d\mathbf{x}d\mathbf{u}_1d\mathbf{u}_2\ldots d\mathbf{u}_{i-1}d\mathbf{u}_{i+1}\ldots d\mathbf{u}_n\int_{\mathbf{u}_i}\nabla_{\mathbf{K}_i}\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})d\mathbf{u}_i\\
&=0.
\end{aligned}
$$

where in the last equation, we have used $\mathbb{Q}^j_{\mathbf{K}}(\mathbf{x}_{\mathcal{N}_j^\kappa},\mathbf{u}_{\mathcal{N}_j^\kappa})$ does not depend on $\mathbf{u}_i$ as $i\notin\mathcal{N}_j^\kappa$, and $\int_{\mathbf{u}_i}\nabla_{\mathbf{K}_i}\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})d\mathbf{u}_i = \nabla_{\mathbf{K}_i}\int_{\mathbf{u}_i}\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})d\mathbf{u}_i = \nabla_{\mathbf{K}_i}1 = 0$. Now we have shown that $E_2 = 0$, then we further bound E_1 as follows,

$$
\begin{aligned}
\|\widehat{\mathbf{h}}_i(\mathbf{K}) - \nabla_{\mathbf{K}_i}C(\mathbf{K})\| = \|E_1\|
&\le \frac{1}{n}\mathbb{E}_{\mathbf{x}\sim\nu_{\mathbf{K}},\mathbf{u}\sim\pi_{\mathbf{K}}(\cdot|x)}\left|\sum_{j\in\mathcal{N}}\mathbb{Q}^i_{\mathbf{K},\kappa}(\mathbf{x}_{\mathcal{N}_i^\kappa},\mathbf{u}_{\mathcal{N}_i^\kappa}) - \sum_{j\in\mathcal{N}}Q^i_{\mathbf{K}}(\mathbf{x},\mathbf{u})\right|\left\|\nabla_{\mathbf{K}_i}\log\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})\right\|\\
&\le \rho^{\kappa+1}\mathbb{E}_{\mathbf{x}\sim\nu_{\mathbf{K}},\mathbf{u}\sim\pi_{\mathbf{K}}(\cdot|\mathbf{x})}C(\mathbf{x},\mathbf{u})\left\|\nabla_{\mathbf{K}_i}\log\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})\right\|\\
&\le CL_i\rho^{\kappa+1},
\end{aligned}
$$

where the last inequality results from Lemma IV.2, (a). We additionally have,

$$\mathbb{E}_{\mathbf{x}\sim\nu_{\mathbf{K}},\mathbf{u}\sim\pi_{\mathbf{K}}(\cdot|\mathbf{x})}\left\|\nabla_{\mathbf{K}_i}\log\pi_{\mathbf{K}_i}(\mathbf{u}_i|\mathbf{x}_{\mathcal{N}_i^r})\right\| = \mathbb{E}_{\mathbf{x}\sim\nu_{\mathbf{K}},\mathbf{u}\sim\pi_{\mathbf{K}}(\cdot|\mathbf{x})}\left\|-\sigma_0^{-2}(\mathbf{u}_i+\mathbf{K}_i\mathbf{x}_{\mathcal{N}_i^r})\mathbf{x}_{\mathcal{N}_i^r}^{\mathrm{T}}\right\| = \sigma_0^{-2}\mathbb{E}\|\varkappa_i(\nu_{\mathbf{K}})_{\mathcal{N}_i^r}^{\mathrm{T}}\| := L_i,$$

where let $\mathbf{u}_i = -\mathbf{K}_i\mathbf{x}_{\mathcal{N}_i^r} + \varkappa_i$, and we have $\nu_{\mathbf{K}} \sim N(\mathbf{0}, \Xi_{\mathbf{K}})$ and $\varkappa_i \sim N(\mathbf{0}, \boldsymbol{I}_{d_u^i})$.

□

IV. Detailed Discussion about the Localized Gradient Approximation

We further define $\mathbf{P}_{\mathbf{K}}^i$ as the unique positive definition solution to the Bellman equation from a local perspective.

$$\mathbf{P}_{\mathbf{K}}^i = (\mathbf{Q}_i + \mathbf{K}^{\mathrm{T}}\mathbf{R}_i\mathbf{K}) + (\mathbf{A}-\mathbf{B}\mathbf{K})^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i(\mathbf{A}-\mathbf{B}\mathbf{K}). \tag{28}$$

We can further define a localized Q function,

$$Q_{\mathbf{K}}^i = \sum_{t=0}^{\infty}\mathbb{E}\big[c_i(t) - C_i(\mathbf{K})\,\big|\,\mathbf{x}(0)=\mathbf{x},\mathbf{u}(0)=\mathbf{u},\mathbf{u}(t)\sim\pi_{\mathbf{K}}(\cdot|\mathbf{x}(t))\big].$$

It is obvious that $Q_{\mathbf{K}}(\mathbf{x},\mathbf{u}) = \frac{1}{n}\sum_{i=1}^n Q_{\mathbf{K}}^i(\mathbf{x},\mathbf{u})$. We represent $Q_{\mathbf{K}}^i(\mathbf{x},\mathbf{u})$ in a quadratic form as well

$$\begin{aligned}Q_{\mathbf{K}}^i(\mathbf{x},\mathbf{u}) =& \mathbf{x}^{\mathrm{T}}(\mathbf{Q}_i + \mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A})\mathbf{x} + \mathbf{x}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B})\mathbf{u}\\ &+ \mathbf{u}^{\mathrm{T}}(\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A})\mathbf{x} + \mathbf{u}^{\mathrm{T}}(\mathbf{R}_i + \mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B})\mathbf{u}\\ &- \mathrm{Tr}(\mathbf{P}_{\mathbf{K}}^i\Xi_{\mathbf{K}}) - {\sigma_0}^2\,\mathrm{Tr}(\mathbf{R} + \mathbf{P}_{\mathbf{K}}^i\mathbf{B}\mathbf{B}^{\mathrm{T}}).\end{aligned}$$

Before proving Lemma IV.16 and Lemma IV.17, we present another lemma as follow,

Lemma 14. If $\mathbf{P}_{\mathbf{K}}^i$ is $(C_{\mathbf{P}_{\mathbf{K}}^i},\rho)$ L-SED, then we have (C,ρ) exponential decay property holds.

Proof. With the explicit representation as shown in Lemma 6, we have,

$$\begin{aligned}\left|Q_{\mathbf{K}}^i(\mathbf{x},\mathbf{u}) - Q_{\mathbf{K}}^i(\mathbf{x}',\mathbf{u}')\right| \le& \left|\mathbf{x}^{\mathrm{T}}(\mathbf{Q}_i + \mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A})\mathbf{x} - \mathbf{x}'^{\mathrm{T}}(\mathbf{Q}_i + \mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A})\mathbf{x}'\right| + \left|\mathbf{x}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B})\mathbf{u} - \mathbf{x}'^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B})\mathbf{u}'\right|\\ &+ \left|\mathbf{u}^{\mathrm{T}}(\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A})\mathbf{x} - \mathbf{u}'^{\mathrm{T}}(\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A})\mathbf{x}'\right| + \left|\mathbf{u}^{\mathrm{T}}(\mathbf{R}_i + \mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B})\mathbf{u} - \mathbf{u}'^{\mathrm{T}}(\mathbf{R}_i + \mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B})\mathbf{u}'\right|\\ =& \left|\mathbf{x}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A})\mathbf{x} - \mathbf{x}'^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A})\mathbf{x}'\right| + \left|\mathbf{x}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B})\mathbf{u} - \mathbf{x}'^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B})\mathbf{u}'\right|\\ &+ \left|\mathbf{u}^{\mathrm{T}}(\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A})\mathbf{x} - \mathbf{u}'^{\mathrm{T}}(\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A})\mathbf{x}'\right| + \left|\mathbf{u}^{\mathrm{T}}(\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B})\mathbf{u} - \mathbf{u}'^{\mathrm{T}}(\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B})\mathbf{u}'\right|\\ :=& A_1 + A_2 + A_3 + A_4,\end{aligned}$$

where in the first equation, we have $\mathbf{x}^{\mathrm{T}}\mathbf{Q}_i\mathbf{x} = \mathbf{x}'^{\mathrm{T}}\mathbf{Q}_i\mathbf{x}'$ and $\mathbf{u}^{\mathrm{T}}\mathbf{R}_i\mathbf{u} = \mathbf{u}'^{\mathrm{T}}\mathbf{R}_i\mathbf{u}'$ since $[\mathbf{Q}_i]_{mk} = 0$ for any $m \notin \mathcal{N}_i^2$ or $k \notin \mathcal{N}_i^2$, and $(\mathbf{R}_i)_{mk} = 0$ for any $m \notin \mathcal{N}_i$ or $k \notin \mathcal{N}_i$. Note that this proof assumes $\kappa \ge 2$ for clarity.

To analyze four terms in the last equation, we must figure out the structure of four core matrices $\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}$, $\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B}$, $\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}$ and $\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B}$. For the (m,k)-th sub-matrix of $\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}$, we have,

$$\begin{aligned}\|[\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}]_{mk}\| =& \left\|\sum_{a=1}^n\sum_{b=1}^n[\mathbf{A}^{\mathrm{T}}]_{ma}[\mathbf{P}_{\mathbf{K}}^i]_{ab}[\mathbf{A}]_{bk}\right\|\\ \le& \sum_{a=1}^n\sum_{b=1}^n\left\|[\mathbf{A}^{\mathrm{T}}]_{ma}\right\|\cdot\left\|[\mathbf{P}_{\mathbf{K}}^i]_{ab}\right\|\cdot\left\|[\mathbf{A}]_{bk}\right\|\\ =& \sum_{a\in\mathcal{N}_m^2}\sum_{b\in\mathcal{N}_k^2}\left\|[\mathbf{A}^{\mathrm{T}}]_{ma}\right\|\cdot\left\|[\mathbf{P}_{\mathbf{K}}^i]_{ab}\right\|\cdot\left\|[\mathbf{A}]_{bk}\right\|\\ =& \sum_{a\in\mathcal{N}_m^2}\sum_{b\in\mathcal{N}_k^2}\left\|[\mathbf{A}^{\mathrm{T}}]_{ma}\right\|\cdot\left\|[\mathbf{A}]_{bk}\right\|\cdot C_{\mathbf{P}_{\mathbf{K}}^i}\rho^{\mathrm{dist}(i,a)+\mathrm{dist}(b,i)}\\ \le& \frac{(N_{\mathcal{G}}^2)^2\overline{[\mathbf{A}]}^2 C_{\mathbf{P}_{\mathbf{K}}^i}}{\rho^4}\rho^{\mathrm{dist}(i,m)+\mathrm{dist}(i,k)},\end{aligned}$$

where in the last inequality, for all $a \in \mathcal{N}_m^2$ and $b \in \mathcal{N}_k^2$, it holds that $\rho^{\mathrm{dist}(i,a)+\mathrm{dist}(i,b)} \leq \rho^{\mathrm{dist}(i,m)-2+\mathrm{dist}(i,k)-2}$. $N_{\mathcal{G}}^2$ is the number of agents in the biggest 2-hop neighborhood. In another word, $\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}$ satisfies $(C_{\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}}, \rho)$ L-SED, where $C_{\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}} = \frac{(N_{\mathcal{G}}^2)^2\overline{[\mathbf{A}]}^2 C_{\mathbf{P}_{\mathbf{K}}^i}}{\rho^4}$. We can also use Lemma 3 to get the same result. Similarly, we have,

$$\|(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B})_{mk}\|_\infty \leq \frac{N_{\mathcal{G}}^2(\overline{[\mathbf{A}]})(\overline{[\mathbf{B}]})C_{\mathbf{P}_{\mathbf{K}}^i}}{\rho^2}\rho^{\mathrm{dist}(m,i)+\mathrm{dist}(k,i)} := C_{\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B}} \cdot \rho^{\mathrm{dist}(m,i)+\mathrm{dist}(k,i)}$$
$$\|(\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A})_{mk}\|_\infty \leq \frac{N_{\mathcal{G}}^2(\overline{[\mathbf{A}]})(\overline{[\mathbf{B}]})C_{\mathbf{P}_{\mathbf{K}}^i}}{\rho^2}\rho^{\mathrm{dist}(m,i)+\mathrm{dist}(k,i)} := C_{\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B}} \cdot \rho^{\mathrm{dist}(m,i)+\mathrm{dist}(k,i)}$$
$$\|(\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B})_{mk}\|_\infty \leq \overline{[\mathbf{B}]}^2 C_{\mathbf{P}_{\mathbf{K}}^i}\rho^{\mathrm{dist}(m,i)+\mathrm{dist}(k,i)} := C_{\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B}} \cdot \rho^{\mathrm{dist}(m,i)+\mathrm{dist}(k,i)}.$$

So that,

$$\begin{aligned}
A_1 \leq & \left|\mathbf{x}_{\mathcal{N}_i^\kappa}^{\mathrm{T}}[\mathbf{Q}_i + \mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}]_{\mathcal{N}_i^\kappa \times \mathcal{N}_{-i}^\kappa}(\mathbf{x}_{\mathcal{N}_{-i}^\kappa} - \mathbf{x}'_{\mathcal{N}_{-i}^\kappa})\right| + \left|(\mathbf{x}_{\mathcal{N}_{-i}^\kappa} - \mathbf{x}'_{\mathcal{N}_{-i}^\kappa})^{\mathrm{T}}[\mathbf{Q}_i + \mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}]_{\mathcal{N}_{-i}^\kappa \times \mathcal{N}_i^\kappa}\mathbf{x}_{\mathcal{N}_i^\kappa}\right| \\
& + \left|\mathbf{x}_{\mathcal{N}_{-i}^\kappa}^{\mathrm{T}}[\mathbf{Q}_i + \mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}]_{\mathcal{N}_{-i}^\kappa \times \mathcal{N}_{-i}^\kappa}\mathbf{x}_{\mathcal{N}_{-i}^\kappa}\right| + \left|\mathbf{x}'^{\mathrm{T}}_{\mathcal{N}_{-i}^\kappa}[\mathbf{Q}_i + \mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}]_{\mathcal{N}_{-i}^\kappa \times \mathcal{N}_{-i}^\kappa}\mathbf{x}'_{\mathcal{N}_{-i}^\kappa}\right| \\
= & \left|\mathbf{x}_{\mathcal{N}_i^\kappa}^{\mathrm{T}}[\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}]_{\mathcal{N}_i^\kappa \times \mathcal{N}_{-i}^\kappa}(\mathbf{x}_{\mathcal{N}_{-i}^\kappa} - \mathbf{x}'_{\mathcal{N}_{-i}^\kappa})\right| + \left|(\mathbf{x}_{\mathcal{N}_{-i}^\kappa} - \mathbf{x}'_{\mathcal{N}_{-i}^\kappa})^{\mathrm{T}}[\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}]_{\mathcal{N}_{-i}^\kappa \times \mathcal{N}_i^\kappa}\mathbf{x}_{\mathcal{N}_i^\kappa}\right| \\
& + \left|\mathbf{x}_{\mathcal{N}_{-i}^\kappa}^{\mathrm{T}}[\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}]_{\mathcal{N}_{-i}^\kappa \times \mathcal{N}_{-i}^\kappa}\mathbf{x}_{\mathcal{N}_{-i}^\kappa}\right| + \left|\mathbf{x}'^{\mathrm{T}}_{\mathcal{N}_{-i}^\kappa}[\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}]_{\mathcal{N}_{-i}^\kappa \times \mathcal{N}_{-i}^\kappa}\mathbf{x}'_{\mathcal{N}_{-i}^\kappa}\right| \\
\leq & 4\|\mathbf{x}\|_\infty^2|\mathcal{N}_i^\kappa||\mathcal{N}_{-i}^\kappa|C_{\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}}\rho^\kappa + 2\|\mathbf{x}\|_\infty^2|\mathcal{N}_{-i}^\kappa|^2 C_{\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{A}}\rho^{2\kappa}.
\end{aligned}$$

where in the last inequality, we have that $\forall(m,n) \in \mathcal{N}_i^\kappa \times \mathcal{N}_{-i}^\kappa$, $\rho^{\mathrm{dist}(m,i)+\mathrm{dist}(n,i)} \leq \rho^\kappa$, and $\forall(m,n) \in \mathcal{N}_{-i}^\kappa \times \mathcal{N}_{-i}^\kappa$, $\rho^{\mathrm{dist}(m,i)+\mathrm{dist}(n,i)} \leq \rho^{2\kappa}$.

Similarly, we have

$$\begin{aligned}
A_2 &\leq 4\|\mathbf{x}\|_\infty\|\mathbf{u}\|_\infty|\mathcal{N}_i^\kappa||\mathcal{N}_{-i}^\kappa|C_{\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B}}M^\kappa + 2\|\mathbf{x}\|_\infty\|\mathbf{u}\|_\infty|\mathcal{N}_{-i}^\kappa|^2 C_{\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B}}\rho^{2\kappa}, \\
A_3 &\leq 4\|\mathbf{x}\|_\infty\|\mathbf{u}\|_\infty|\mathcal{N}_i^\kappa||\mathcal{N}_{-i}^\kappa|C_{\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B}}M^\kappa + 2\|\mathbf{x}\|_\infty\|\mathbf{u}\|_\infty|\mathcal{N}_{-i}^\kappa|^2 C_{\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B}}\rho^{2\kappa}, \\
A_4 &\leq 4\|\mathbf{u}\|_\infty|\mathcal{N}_i^\kappa||\mathcal{N}_{-i}^\kappa|C_{\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B}}M^\kappa + 2\|\mathbf{u}\|_\infty^2|\mathcal{N}_{-i}^\kappa|^2 C_{\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}^i\mathbf{B}}\rho^{2\kappa}.
\end{aligned}$$

By summing up the upper-bound of A_1, A_2, A_3, A_4, we have:

$$|Q_{\mathbf{K}}^i(\mathbf{x},\mathbf{u}) - Q_{\mathbf{K}}^i(\mathbf{x}',\mathbf{u}')| \leq C_{Q_{\mathbf{K}}^i}(\mathbf{x},\mathbf{u})\rho^{\kappa+1},$$

where the constant $C_{Q_{\mathbf{K}}^i}$ depends on $C_{\mathbf{P}_{\mathbf{K}}^i}$, the system parameter $\mathbf{A}$, $\mathbf{B}$, $\mathbf{Q}$, $\mathbf{R}$, $\|\mathbf{x}\|_\infty$ and $\|\mathbf{u}\|_\infty$.

Now we are prepared to illustrate that Lemma IV.16 and Lemma IV.17 can both leads to the establishment of exponential decay property.

□

A. Proof of Lemma IV.16

Proof. We set $\mathbf{A}_{\mathbf{K}} = \mathbf{A} - \mathbf{B}\mathbf{K}$ for simplicity. For any node j, suppose that i, i' are κ, $\kappa+1$-hop neighbor of j respectively, for any K in the gradient descent process,

$$\frac{\|(\mathbf{A}_{\mathbf{K}}{}^t)_{i'j}\|}{\|(\mathbf{A}_{\mathbf{K}}{}^t)_{ij}\|} \leq M < 1. \tag{29}$$

First, we show that with the condition above, L-SED property of $\mathbf{P}_{\mathbf{K}}$ is guaranteed. As $\rho(\mathbf{A}_{\mathbf{K}}) < 1$, there exists a constant $H > 0$ and $\alpha \in (0,1)$ that $\|(\mathbf{A}_{\mathbf{K}})^t\| < H\alpha^t$ as shown in 1.

For any i, j:

$$\|[\mathbf{A}_{\mathbf{K}}{}^t]_{ij}\| \leq \|[\mathbf{A}_{\mathbf{K}}{}^t]_{ii}\|\rho^{\mathrm{dist}(i,j)} \leq \sqrt{d}\|\mathbf{A}_{\mathbf{K}}{}^t\|\rho^{\mathrm{dist}(i,j)} \leq \sqrt{d}H\alpha^t\rho^{\mathrm{dist}(i,j)} \leq \sqrt{d}H\alpha^t\rho^{\mathrm{dist}(i,j)}. \tag{30}$$

We have

$$\mathbf{P}_{\mathbf{K}}^i = \mathcal{T}_{\mathbf{K}}^{\mathrm{T}}(\mathbf{Q}_i + \mathbf{K}^{\mathrm{T}}\mathbf{R}_i\mathbf{K}) = \sum_{t\geq 0}[(\mathbf{A}-\mathbf{B}\mathbf{K})^t]^{\mathrm{T}}(\mathbf{Q}_i + \mathbf{K}^{\mathrm{T}}\mathbf{R}_i\mathbf{K})(\mathbf{A}-\mathbf{B}\mathbf{K})^{\mathrm{T}} := \sum_{t\geq 0}(\mathbf{A}_{\mathbf{K}}{}^t)^{\mathrm{T}}\mathbf{\Omega}_i\mathbf{A}_{\mathbf{K}}{}^{\mathrm{T}},$$

where in the last equation we use $\mathbf{\Omega}_i$ to denote $\mathbf{Q}_i + \mathbf{K}^{\mathrm{T}}\mathbf{R}_i\mathbf{K}$ for notation simplicity. Thus in the $\mathbf{P}_{\mathbf{K}}^i$, the norm of its $(m.k)$-th sub-matrix can be bounded as,

$$
\begin{aligned}
\|[\mathbf{P}_{\mathbf{K}}^i]_{mk}\| &= \left\| \sum_{t\geq 0}\sum_{a=1}^{n}\sum_{b=1}^{n} [({\mathbf{A}_{\mathbf{K}}}^t)^{\mathrm{T}}]_{ma}[\mathbf{\Omega}_i]_{ab}[{\mathbf{A}_{\mathbf{K}}}^t]_{bk} \right\| \\
&= \left\| \sum_{t\geq 0}\sum_{a\in\mathcal{N}_i^{2r}}\sum_{b\in\mathcal{N}_i^{2r}} [({\mathbf{A}_{\mathbf{K}}}^t)^{\mathrm{T}}]_{ma}[\mathbf{\Omega}_i]_{ab}[{\mathbf{A}_{\mathbf{K}}}^t]_{bk} \right\| \\
&\leq \sum_{t\geq 0}\sum_{a\in\mathcal{N}_i^{2r}}\sum_{b\in\mathcal{N}_i^{2r}} \left\|[({\mathbf{A}_{\mathbf{K}}}^t)^{\mathrm{T}}]_{ma}\right\| \cdot \left\|[\mathbf{\Omega}_i]_{ab}\right\| \cdot \left\|[{\mathbf{A}_{\mathbf{K}}}^t]_{bk}\right\| \\
&\leq \sum_{t\geq 0}\sum_{a\in\mathcal{N}_i^{2r}}\sum_{b\in\mathcal{N}_i^{2r}} d\widehat{\mathbf{\Omega}}_i H^2\alpha^{2t}\rho^{\mathrm{dist}(m,a)+\mathrm{dist}(b,k)} \\
&\leq \sum_{t\geq 0}\sum_{a\in\mathcal{N}_i^{2r}}\sum_{b\in\mathcal{N}_i^{2r}} d\widehat{\mathbf{\Omega}}_i H^2\alpha^{2t}\rho^{\mathrm{dist}(m,i)+\mathrm{dist}(i,k)-4r} \\
&= \sum_{t\geq 0} |\mathcal{N}_i^{2r}|^2 d\widehat{\mathbf{\Omega}}_i H^2\alpha^{2t}\rho^{-4r}\rho^{\mathrm{dist}(m,i)+\mathrm{dist}(i,k)} \\
&= \frac{|\mathcal{N}_i^{2r}|^2 dH^2\widehat{\mathbf{\Omega}}_i\rho^{-4r}}{1-\alpha^2}\rho^{\mathrm{dist}(m,i)+\mathrm{dist}(i,k)} \\
&:= C_{\mathbf{P}_{\mathbf{K}}^i}\rho^{\mathrm{dist}(m,i)+\mathrm{dist}(i,k)},
\end{aligned}
$$

where we have $\mathbf{\Omega}_i = \mathbf{Q}_i + \mathbf{K}^T\mathbf{R}_i\mathbf{K}$ and $\mathbf{\Omega}_1 \in \mathcal{M}^2$. Moreover, in the third inequality, we have that, for any agent $m \notin \mathcal{N}_i^{2r}$ and any agent $a \in \mathcal{N}_i^{2r}$, it holds that $\rho^{\mathrm{dist}(m,a)} \leq \rho^{\mathrm{dist}(m,i)-2r}$.

So $\mathbf{P}_{\mathbf{K}}^i$ is $(C_{\mathbf{P}_{\mathbf{K}}^i}, M)$ L-SED. As $\mathbf{P}_{\mathbf{K}} = \frac{1}{n}\sum_{i=1}^n \mathbf{P}_{\mathbf{K}}^i$, we have $\|[\mathbf{P}_{\mathbf{K}}]_{mk}\| \leq \frac{1}{n}\sum_{i=1}^n C_{\mathbf{P}_{\mathbf{K}}^i}$. In another word, $\mathbf{P}_{\mathbf{K}}$ is $(C_{\mathbf{P}_{\mathbf{K}}}, M)$ SED, where $C_{\mathbf{P}_{\mathbf{K}}} = \frac{1}{n}\sum_{i=1}^n \frac{|\mathcal{N}_i^{2r}|^2 dH^2\widehat{\mathbf{\Omega}}_i\rho^{-4r}}{1-\alpha^2}$. □

B. Proof of Lemma IV.17

Proof. We have

$$
\begin{aligned}
\|[(\mathbf{A}_{\mathbf{K}})_{ij}^t]\| &= \left\| \sum_{n_1,n_2,\ldots,n_{t-1}\,\middle|\, i\to n_1\to n_2\to\ldots\to n_{t-1}\to j\in\mathcal{W}_{i\to j}^t(r)} \mathbf{A}_{\mathbf{K}i,n_1}\mathbf{A}_{\mathbf{K}n_1,n_2}\ldots\mathbf{A}_{\mathbf{K}n_{t-2},n_{t-1}}\mathbf{A}_{\mathbf{K}n_{t-1},j} \right\| \\
&\leq \sum_{n_1,n_2,\ldots,n_{t-1}\,|\, i\to n_1\to n_2\to\ldots\to n_{t-1}\to j\in\mathcal{W}_{i\to j}^t(r)} \left\|\mathbf{A}_{\mathbf{K}i,n_1}\mathbf{A}_{\mathbf{K}n_1,n_2}\ldots\mathbf{A}_{\mathbf{K}n_{t-2},n_{t-1}}\mathbf{A}_{\mathbf{K}n_{t-1},j}\right\| \\
&\leq \sum_{n_1,n_2,\ldots,n_{t-1}\,|\, i\to n_1\to n_2\to\ldots\to n_{t-1}\to j\in\mathcal{W}_{i\to j}^t(r)} (\max_{a_1,a_2\in\{1,2,3,\ldots,n\}} \|\mathbf{A}_{\mathbf{K}a_1,a_2}\|)^t \\
&\leq (\max_{a_1,a_2\in\{1,2,3,\ldots,n\}} \|\mathbf{A}_{\mathbf{K}a_1,a_2}\|)^t |\mathcal{W}_{i\to j}^t(r)| \\
&\leq C\rho^{\kappa}(\mathcal{D}\overline{[\mathbf{A}_{\mathbf{K}}]})^t,
\end{aligned}
$$

where we obtain the same form as (30). According to the derivation similar to the proof in Section IV-A, we can prove the SED of $\mathbf{P}_{\mathbf{K}}$ and the L-SED of $\mathbf{P}_{\mathbf{K}}^i$. □

C. The Spatial Exponential Decay Property of $\nabla_{\mathbf{K}} C(\mathbf{K})$

Lemma 15. If Lemma IV.16 or Lemma IV.17 is satisfied, the matrix $\nabla_{\mathbf{K}} C(\mathbf{K})$ enjoys a $(C_{\nabla_{\mathbf{K}}}, \rho)$ spatially exponential decaying (SED),

$$
C_{\nabla_{\mathbf{K}}} = 2n\left[(n^3 C_{\mathbf{P}_{\mathbf{K}}}\overline{[\mathbf{B}]}^2 + n\overline{[\mathbf{R}]})\overline{[\mathbf{K}]} + n^2 C_{\mathbf{P}_{\mathbf{K}}}\overline{[\mathbf{B}]}\cdot\overline{[\mathbf{A}]}\right] C_{\mathbf{\Xi}_{\mathbf{K}}},
$$

where $C_{\mathbf{P}_{\mathbf{K}}}$ is defined in Lemma IV-A and,

$$
C(\mathbf{\Xi}_{\mathbf{K}}) = \frac{ndH^2(\sigma^2 + \sigma_0^2\|\mathbf{B}\|^2)}{1-\alpha^2}.
$$

Proof. Revisiting Lemma 6, we have

$$\nabla_{\mathbf{K}} C(\mathbf{K}) = 2[(\mathbf{R} + \mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{B})\mathbf{K} - \mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathbf{K}}\mathbf{A}]\mathbf{\Xi}_{\mathbf{K}}.$$

$\mathbf{P}_{\mathbf{K}}$ satisfies $(C_{\mathbf{P}_{\mathbf{K}}}, \rho)$-SED as shown in Section IV-A. Then we analyze the term $\mathbf{\Xi}_{\mathbf{K}}$.

$$\mathbf{\Xi}_{\mathbf{K}} = \mathcal{T}_{\mathbf{K}}(\mathbf{\Psi}) =:= \sum_{t\geq 0} [\mathbf{A}_{\mathbf{K}}^t]^{\mathrm{T}}(\sigma^2 \boldsymbol{I} + \sigma_0^2 \mathbf{B}\mathbf{B}^{\mathrm{T}})\mathbf{A}_{\mathbf{K}}^t.$$

For any agent m and k, we additionally analyze the bound of $\|[\mathbf{\Xi}_{\mathbf{K}}]_{mk}\|$.

$$\begin{aligned}
\|[\mathbf{\Xi}_{\mathbf{K}}]_{mk}\| &= \left\| \sum_{t\geq 0}\sum_{a=1}^{n}\sum_{b=1}^{n} [(\mathbf{A}_{\mathbf{K}}{}^t)^{\mathrm{T}}]_{ma}[\sigma^2 \boldsymbol{I} + \sigma_0^2\mathbf{B}\mathbf{B}^{\mathrm{T}}]_{ab}[\mathbf{A}_{\mathbf{K}}{}^t]_{bk} \right\| \\
&= \left\| \sum_{t\geq 0}\sum_{a=1}^{n} [(\mathbf{A}_{\mathbf{K}}{}^t)^{\mathrm{T}}]_{ma}[\sigma^2 \boldsymbol{I} + \sigma_0^2\mathbf{B}\mathbf{B}^{\mathrm{T}}]_{aa}[\mathbf{A}_{\mathbf{K}}{}^t]_{ak} \right\| \\
&\leq \sum_{t\geq 0}\sum_{a=1}^{n} dH^2\alpha^{2t}\rho^{\mathrm{dist}(m,a)+\mathrm{dist}(a,k)}(\sigma^2 + \sigma_0^2\|\mathbf{B}\|^2) \\
&\leq \sum_{t\geq 0}\sum_{a=1}^{n} dH^2\alpha^{2t}\rho^{\mathrm{dist}(m,k)}(\sigma^2 + \sigma_0^2\|\mathbf{B}\|^2) \\
&\leq \frac{ndH^2(\sigma^2 + \sigma_0^2\|\mathbf{B}\|^2)}{1-\alpha^2}\rho^{\mathrm{dist}(m,k)} \\
&:= C_{\mathbf{\Xi}_{\mathbf{K}}}\rho^{\mathrm{dist}(m,k)}.
\end{aligned}$$

Note that $\mathbf{R}, \mathbf{B} \in \mathcal{M}^1$, $\mathbf{A} \in \mathcal{M}^2$, $\mathbf{K} \in \mathcal{M}^r$. Then with the help of Lemma 2, Lemma 3 and Lemma 4, we have $\nabla_{\mathbf{K}} C(\mathbf{K})$ is $(C_{\nabla_{\mathbf{K}}}, \rho)$-SED, where

$$C_{\nabla_{\mathbf{K}}} = 2n\left[(n^3 C_{\mathbf{P}_{\mathbf{K}}}\overline{[\mathbf{B}]}^2 + n\overline{[\mathbf{R}]})\overline{[\mathbf{K}]} + n^2 C_{\mathbf{P}_{\mathbf{K}}}\overline{[\mathbf{B}]}\cdot\overline{[\mathbf{A}]}\right]C_{\mathbf{\Xi}_{\mathbf{K}}}.$$

□

V. Experimental Results

We choose four representative graphs, line, circle, tree, and 4-regular grid, to verify our main theorem (Theorem IV.3). We emphasize that the combination of these graphs can represent a large class of random graphs.

In the line and the circle graph, we set 99 nodes. In the 2-ary tree graph, we set 127 nodes to construct an 8-layer full binary tree. In the 4-regular grid, we set $11 \times 11 = 121$ nodes. A is set to be the scaling adjacency matrix that $\rho(A) < 1$. The cost matrices Q, R are taken to be identity with appropriate dimensions. We set step size η is set to 0.001. The noise covariance matrix $\mathbf{\Psi}$ is set to be $0.5I$. The total iteration T is set to be 4000.

We present the relative cost error $\frac{C(\mathbf{K}(T))-C(\mathbf{K}^*)}{C(\mathbf{K}^*)}$ with different control dependence range r in Figure 2, $r \in [2, 3, 5, 10, 20]$. Note that we set κ to be the maximum so that every agent i has access to the exact gradient $\nabla_{\mathbf{K}_i} C(\mathbf{K})$. $r = 2$ means that the dynamic of every agent only depends on its 2-hop neighbors. On the contrary, $r = 20$ means that every agent can affect any other agents in the graph, and the situation degenerates to the centralized LQR. Thus the controller can converge to the optima without degradation.

We present the relative cost error $\frac{C(\mathbf{K}(T))-C(\mathbf{K}^*)}{C(\mathbf{K}^*)}$ with different communication range limits κ in Figure 3, $\kappa \in [2, 3, 5, 10, 20]$. Note that r is set to maximum, meaning there will be no gradient truncation. C is set to be 0.1, ρ is set to be 0.9, and the approximation error is added on $\nabla_{\mathbf{K}_i} C(\mathbf{K})$ following a Gaussian distribution. Note that we simulate the approximation error here. Scalable model-free algorithms with Monte-Carlo and Actor-Critic will be our next work.

The results verify our main conclusion that, as r and κ increase, the performance gap between $\mathbf{K}(T)$ and $\mathbf{K}^*$ decreases exponentially. With a finite communication range κ, it is possible to design a scalable and decentralized gradient descent algorithm to obtain a near-optimal controller.

VI. Examples and Numerical Results for Exponential Decay Property

We give some typical examples and numerical results for Lemma IV.17. Note that it requires that $|\mathcal{W}_{i\to j}^t(r)| \leq C\mathcal{D}^t\rho^\kappa$.

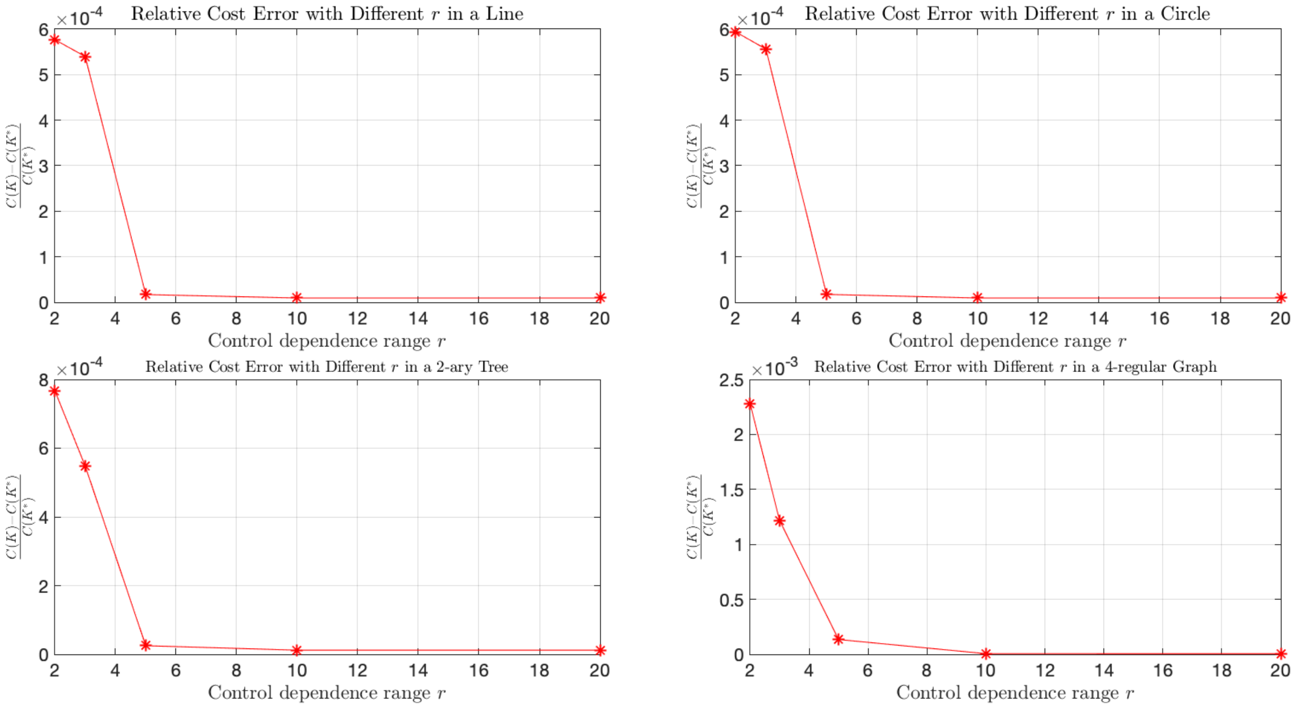


Fig. 2: Relative performance gap with different control dependence range r.

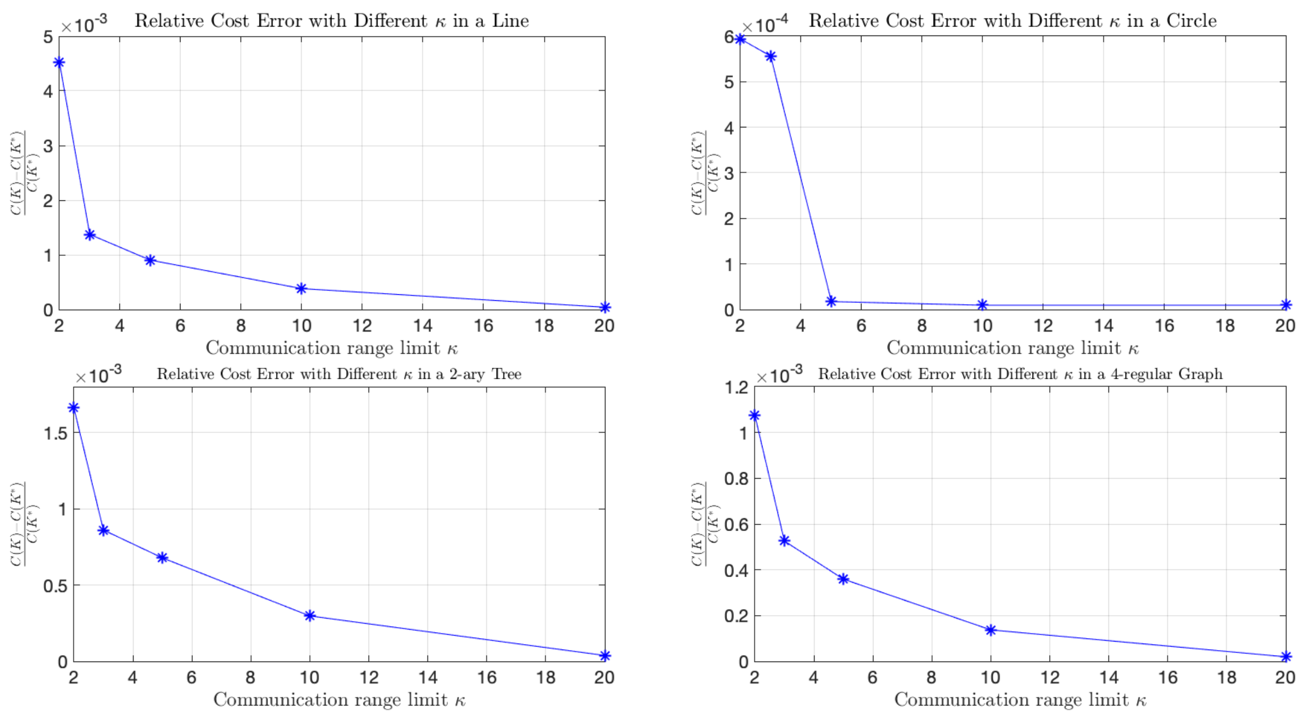


Fig. 3: Relative performance gap with different communication range limit κ.

A. Line

In a line, a walk of length t can be divided into several moves: going to the left, going to the right, and staying still. We assume that in the t-step move, there are x steps to the right, y steps to the left, and z steps staying still. We have $x - y = \kappa$ and $x + y + z = t$. Constraints are that all x, y, and z must be at least 0. So that in this network, if $t \gg \kappa$, we have

$$\mathcal{W}^t_{i\to j}(r=1) = \binom{t}{\kappa} + \binom{t}{\kappa+1}\binom{t-\kappa-1}{1} + \ldots + \binom{t}{\frac{t+\kappa}{2}}.$$

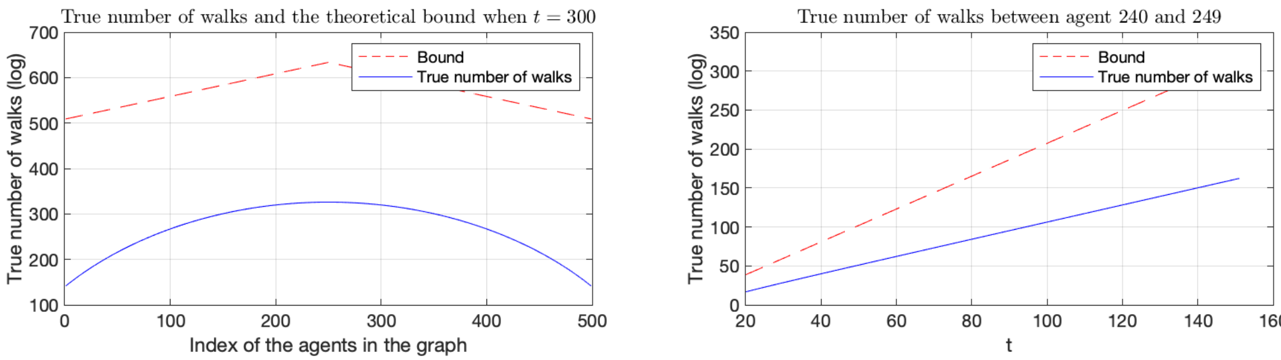


Fig. 4: The number of the true walks $|\mathcal{W}_{i\to j}|$ and the theoretical bound in a 300-node line. In the left, we fix $t = 300$ and vary κ. In the right, We fix $\kappa = 9$ and vary t.

We consider a single term, note that $t \gg \kappa$.

$$\binom{t}{\kappa+x}\binom{t-\kappa-x}{x} = \frac{t!}{(t-\kappa-x)!(k+\kappa)!}\cdot\frac{(t-\kappa-x)!}{(t-\kappa-2x)!x!} = \frac{t!}{(\kappa+x)!(t-\kappa-2x)!x!} < \frac{e(\frac{t}{2})^t}{(\frac{\kappa+x}{e})^{\kappa+x}(\frac{t-\kappa-2x}{e})^{t-\kappa-2x}(\frac{x}{e})^x}$$
$$= \frac{(\frac{t}{2})^t e^{t+1}}{(\kappa+x)^{\kappa+x}(t-\kappa-2x)^{t-\kappa-2x}x^x} \le \frac{(\frac{t}{2})^t e^{t+1}}{(\kappa+\frac{t}{3})^{\kappa+\frac{t}{3}}(\frac{t}{3}-\kappa)^{\frac{t}{3}-\kappa}(\frac{t}{3})^{\frac{t}{3}}} \le (\frac{3}{2})^t\cdot e^{t+1}.$$

For the fifth inequality, we set $f(x) = (k+\kappa)^{k+\kappa}(t-\kappa-2x)^{t-\kappa-2x}x^x$ and $x \in (0, \frac{t-k}{2})$. Now we prove that when $x = \frac{t}{3}$, $f(x)$ reaches a minimum.

$$g(x) = \log f(x) = (x+\kappa)\log(x+\kappa) + (t-\kappa-2x)\log(t-\kappa-2x) + x\log(x),$$
$$g'(x) = \log\frac{(x+\kappa)x}{(t-\kappa-2x)^2},$$

when $x \in (0, \frac{4t-3\kappa-\sqrt{4t^2-3\kappa^2}}{6})$, $g'(x) < 0$, $g(x)$ and $f(x) \downarrow$, when $x \in (\frac{4t-3\kappa-\sqrt{4t^2-3\kappa^2}}{6}, \frac{t-\kappa}{2})$,$g'(x) > 0$, $g(x)$ and $f(x)$ $\uparrow$. The minimum point of $f(x)$ is $x = \frac{4t-3\kappa-\sqrt{4t^2-3\kappa^2}}{6}$. As $t \gg \kappa$, we can take $x = \frac{t}{3}$.

Combining the results above, we have

$$\mathcal{W}^t_{i\to j}(r=1) \le (\frac{t-\kappa}{2}+1)\cdot\frac{3}{2}^t\cdot e^{t+1} \le e^{\frac{t-\kappa}{2}}\frac{3}{2}^t\cdot e^{t+1} = e\cdot[(\frac{3e}{2})^{\frac{3}{2}}]^t\cdot[e^{-\frac{1}{2}}]^\kappa,$$

which has the same form as Lemma IV.17.

We also do some numerical experiments to verify such a result. We set 499 nodes in a line network and set $t = 300$. We pick agent 249 to show the difference between the true walks number and our bound.

B. Cycle

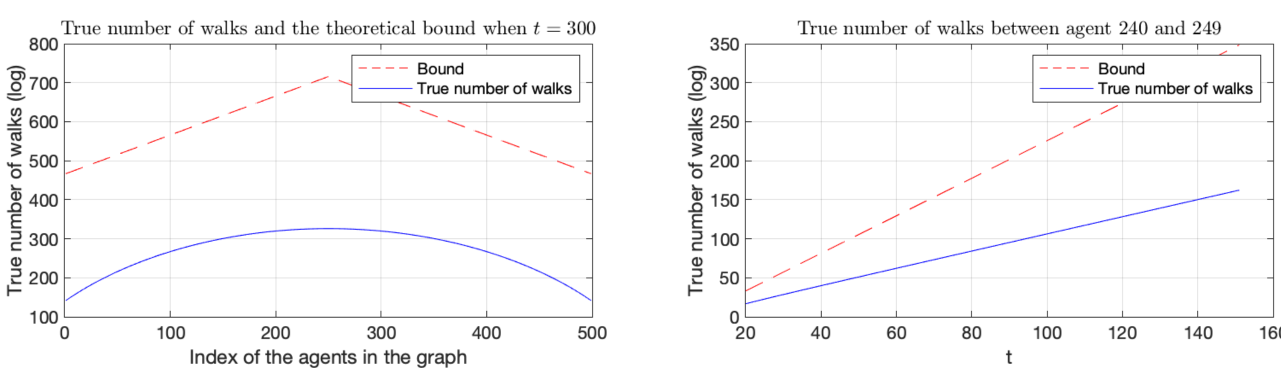
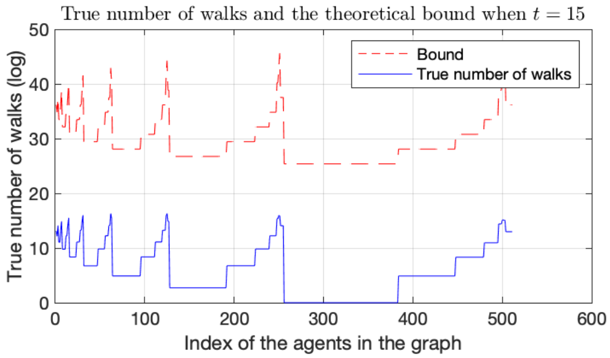


Fig. 5: The number of the true walks $|\mathcal{W}_{i\to j}|$ and the theoretical bound in a 300-node circle. In the left, we fix $t = 300$ and vary κ. In the right, We fix $\kappa = 9$ and vary t.



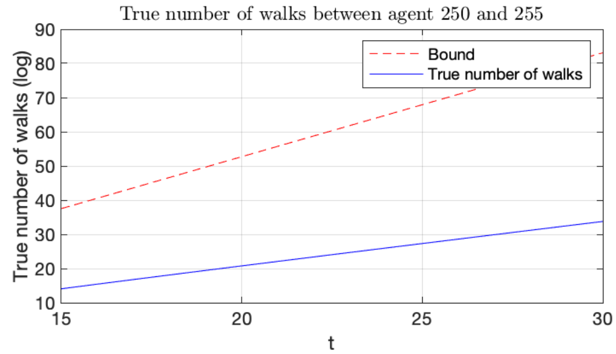


Fig. 6: The number of the true walks $|\mathcal{W}_{i\to j}|$ and the theoretical bound in a 9-layer full binary tree. In the left, we fix $t = 15$ and vary κ. In the right, We fix $\kappa = 5$ and vary t.

In a cycle network, we assume that in the t-step move, there are x steps clockwise, y steps counterclockwise, and z steps staying still. ithout loss of generality, we assume $y - x = c \cdot n + \kappa$, where n represents the total number of agents in the network. We have $x + y + z = t$, $y = x + c \cdot n + \kappa$, where $c \cdot n \in (0, t - \kappa)$. With the assumption that $t \gg \kappa$, we have,

$$
\begin{aligned}
\mathcal{W}^t_{i\to j}(r=1) &= \sum_{x=1}^{\lfloor (t-cn-\kappa)/2 \rfloor} \sum_{c} \binom{t}{x}\binom{t-x}{x+cn+\kappa} = \sum_{x=1}^{\lfloor (t-cn-\kappa)/2 \rfloor} \sum_{c} \frac{t!}{x!(x+cn+\kappa)!(t-2x-cn-\kappa)!} \\
&\leq \sum_{x=1}^{\lfloor (t-cn-\kappa)/2 \rfloor} \sum_{c} \frac{e(\frac{t}{2})^t}{(\frac{x}{e})^x(\frac{x+cn+\kappa}{e})^{x+cn+\kappa}(\frac{t-2x-cn-\kappa}{e})^{t-2x-cn-\kappa}} \\
&= \sum_{x=1}^{\lfloor (t-cn-\kappa)/2 \rfloor} \sum_{c} \frac{e^{t+1}(\frac{t}{2})^t}{(x)^x(x+cn+\kappa)^{x+cn+\kappa}(t-2x-cn-\kappa)^{t-2x-cn-\kappa}} \\
&\approx \sum_{x=1}^{\lfloor (t-cn-\kappa)/2 \rfloor} \sum_{c} \frac{e^{t+1}(\frac{t}{2})^t}{(x)^x(x+cn)^{x+cn}(t-2x-cn)^{t-2x-cn}} \leq \sum_{x=1}^{\lfloor (t-cn-\kappa)/2 \rfloor} \sum_{c} \frac{e^{t+1}(\frac{t}{2})^t}{(x)^x(\frac{t-x}{2})^{t-x}} \\
&\leq \sum_{x=1}^{\lfloor (t-cn-\kappa)/2 \rfloor} \sum_{c} (\frac{3}{2})^t \cdot e^{t+1} \leq \frac{t-\kappa}{2} * \frac{t-\kappa}{n} * (\frac{3}{2})^t \cdot e^{t+1} \leq \frac{e}{2n}(\frac{3}{2}e^2)^t \cdot e^{-\kappa},
\end{aligned}
$$

where in the sixth inequality, $c \cdot n = \frac{t-3x}{2}$. We use $e^x \geq x^2$ in the last step.

C. f-ary Tree

We consider a f-ary tree and assume that agent i and agent j are κ-hop neighbors. We assume that in the t-step move, there are x steps staying still. It is obvious that $0 \leq x \leq t - \kappa$. We have

$$
\begin{aligned}
\mathcal{W}^t_{i\to j} &\leq \sum_x f^{\frac{t-\kappa-x}{2}} \binom{t}{\kappa}\binom{t-\kappa}{x}\binom{t-\kappa-x}{\frac{t-\kappa-x}{2}} \leq \sum_x f^{\frac{t-\kappa-x}{2}} \frac{t!}{\kappa! x! [(\frac{t-\kappa-x}{2})!]^2} \leq \sum_x f^{\frac{t-\kappa-x}{2}} \frac{e^{t+1}(\frac{t}{2})^t}{\kappa^\kappa x^x (\frac{t-\kappa-z}{2})^{t-\kappa-z}} \\
&\leq \sum_x f^{\frac{t-\kappa-x}{2}} \frac{e^{t+1}(\frac{t}{2})^t}{(\frac{t}{4})^t} < \sum_x f^{\frac{t-\kappa}{2}} e^{t+1} 2^t < f^{\frac{t-\kappa}{2}} e^{t+1} 2^t (t-\kappa) < f^{\frac{t-\kappa}{2}} e^{t+1} 2^t e^{t-\kappa-1} \\
&= (2e^2 f^{\frac{1}{2}})^t (e^{-1} f^{-\frac{1}{2}})^\kappa,
\end{aligned}
$$

Here we have to make some explanations for the first inequality. In a d-array tree, there is only ONE walk between any two specified nodes. So we have to pick out these movement out of t steps ($\binom{t}{\kappa}$); then we have x steps for not moving ($\binom{t-\kappa}{x}$); at last, we have $\frac{t-\kappa-x}{2}$ steps to move to the parent node and $\frac{t-\kappa-x}{2}$ steps to move to the children nodes, which has $f^{\frac{t-\kappa-x}{2}} \binom{t-\kappa-x}{\frac{t-\kappa-x}{2}}$ possible cases. Note that such a combination does not guarantee that we find a walk from i to j, so we use $\leq$ in the inequality.

We do simulations on a full binary tree with 511 nodes.

D. 4-regular Grid

We consider a 4-regular grid and assume that agent i and agent j are κ-hop neighbors. In this grid, we assume i and j are κ_x steps horizontally and κ_y steps vertically, so that $\kappa = \kappa_x + \kappa_y$.

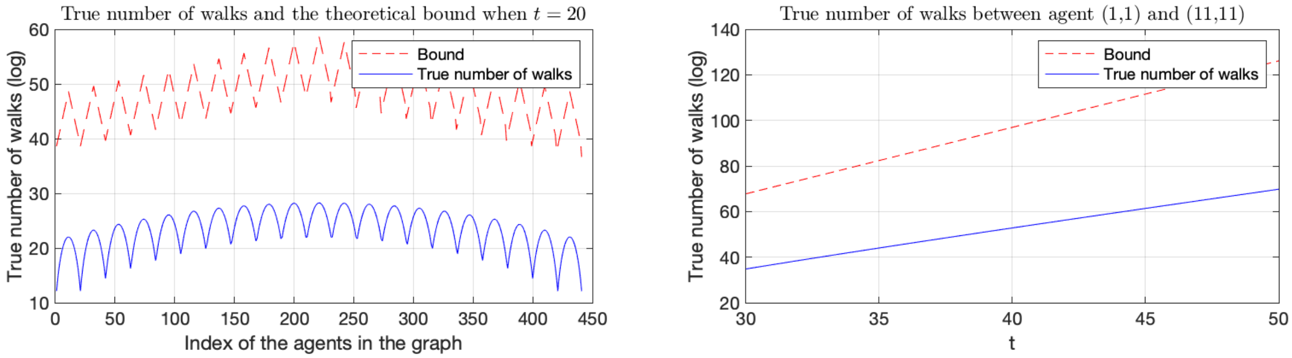


Fig. 7: The number of the true walks $|\mathcal{W}_{i\to j}|$ and the theoretical bound in a 21×21 4-regular grid. In the left, we fix $t = 20$ and vary κ. In the right, We fix $\kappa = 20$ and vary t.

We assume that in the t-step move, there are x steps to the right, y steps to the left, h steps up, g steps down and u steps staying still. We have $y - x = \kappa_x$, $g - u = \kappa_y$, $x + y + h + g + u = t$. So that, $y = x + \kappa_x$, $h = \frac{t-u-\kappa}{2} - x$. We assume $t \gg \kappa, \kappa_x, \kappa_y$. We have $x, y, h, g, u \geq 0$, so that $0 \leq x \leq \frac{t-u-\kappa}{2}$, $0 \leq u \leq t - \kappa$.

$$
\begin{aligned}
\mathcal{W}^t_{i\to j}(r=1) &= \sum_u \sum_x \binom{t}{u}\binom{t-u}{x}\binom{t-u-x}{y}\binom{t-u-x-y}{h} = \sum_u \sum_x \binom{t}{u}\binom{t-u}{x}\binom{t-u-x}{x+\kappa_x}\binom{t-u-2x-\kappa_x}{\frac{t-u-\kappa}{2}-x} \\
&= \sum_u \sum_x \frac{t!}{u!x!(x+\kappa_x)^{x+\kappa_x}(x+\kappa_x)!(\frac{t-u+\kappa}{2} - x + \kappa_x)!(\frac{t-u-\kappa}{2} - x)!} \\
&\approx \sum_u \sum_x \frac{(\frac{t}{2})^t e^{t+1}}{u^u x^{2x}(\frac{t-u}{2} - x)^{2[\frac{t-u}{2} - x]}} \leq \sum_u \sum_x \frac{(\frac{t}{2})^t e^{t+1}}{(\frac{t}{5})^t} \\
&= \sum_u \frac{t-u-\kappa}{2} \cdot \frac{(\frac{t}{2})^t e^{t+1}}{(\frac{t}{5})^t} < \frac{(t-\kappa)^2}{2} \cdot (\frac{5}{2})^t e^{t+1} < \frac{1}{2} e^{t-\kappa} (\frac{5}{2})^t e^{t+1} = \frac{e}{2}(\frac{5e^2}{2})^t (e^{-1})^\kappa,
\end{aligned}
$$

where in the 5'th inequality, it takes $u = x = \frac{t}{5}$. And in the 7'th inequality, we use $e^x > x^2$ when x is large enough.

Remark 5. Note that the tighter theoretical bounds for the number of walks in these graphs could be expected, for we have used a lot of loose inequality scaling technique.

VII. Gradient Descent Related

Lemma 16. [3] $\nabla C(\mathbf{K})$ is Lipschitz continuous with rank L on $S_{C(\mathbf{K}(0))}$.

Lemma 17. Let $L = \sup_{K \in S_{C(\mathbf{K}(0))}} \|\nabla^2 C(\mathbf{K})\|$ and consider the sequence $\{\mathbf{K}_j\}_{j=0}^\infty$ generated by gradient descent with stepsize $\eta \leq \frac{2}{L}$. If $\mathbf{K}(0)$ is stabilizing, then the sequence remains stabilizing.

Proof. The GD update gives $\mathbf{K}(t+1) = \mathbf{K}(t) - \eta \nabla C(\mathbf{K}(t))$, and we have,

$$
C(y) \leq C(x) + \langle \nabla C(x), y - x \rangle + \frac{L}{2}\|y - x\|_2^2.
$$

Put $y = \mathbf{K}(t+1)$, $x = \mathbf{K}(T)$:

$$
\begin{aligned}
C(\mathbf{K}(t+1)) &\leq C(\mathbf{K}(t)) + \langle \nabla C(x), \mathbf{K}(t+1) - \mathbf{K}(T) \rangle + \frac{L}{2}\|\mathbf{K}(t+1) - \mathbf{K}(t)\|^2 \\
&= C(\mathbf{K}(t)) + \langle \nabla C(x), -\eta \nabla C(\mathbf{K}(t)) \rangle + \frac{L}{2}\|\eta \nabla C(\mathbf{K}(t))\|^2 \\
&= C(\mathbf{K}(t)) - (\eta - \frac{L}{2}\eta^2)\|\nabla C(\mathbf{K}(t))\|^2 \leq C(\mathbf{K}(t)).
\end{aligned}
$$

$\square$

VIII. Monte-Carlo version of the algorithm

Algorithm 1 Model-Free Scalable Monte-Carlo-based Policy Gradient (one step)

Input: $\mathbf{K}(t)$ ($\{\mathbf{K}_1(t), \mathbf{K}_2(t), \dots, \mathbf{K}_n(t)\}$), parameter κ, roll out length l_1 and l_2, step size η.

// Approximate $C_i(K)$

for $m = 0, 1, 2, \dots, M_1 - 1$ **do**

Simulate l_1 steps starting with control policy $\mathbf{K}(t)$ in a subroutine, resulting to a sample sequence: $\{(\dot{\mathbf{x}}^m(1), \dot{\mathbf{u}}^m(1)), (\dot{\mathbf{x}}^m(2), \dot{\mathbf{u}}^m(2))), \dots, (\dot{\mathbf{x}}^m(l), \dot{\mathbf{u}}^m(l))\}$. For each agent i,

$$\widehat{C}_i(m) = \frac{1}{l_1} \sum_{\tau=1}^{l_1} \dot{c}_i(\tau) = \frac{1}{l_1} \sum_{\tau=1}^{l_1} [\dot{\mathbf{x}}_{\mathcal{N}_i}^m(\tau)^{\mathrm{T}} \widetilde{\mathbf{Q}}_i \dot{\mathbf{x}}_{\mathcal{N}_i}^m(\tau) + \dot{\mathbf{u}}_i^m(\tau)^{\mathrm{T}} \widetilde{\mathbf{R}}_i \dot{\mathbf{u}}_i^m(\tau)] \tag{31}$$

end for

$$\widehat{C}_i = \frac{1}{M_1} \sum_{m=0}^{M_1-1} \widehat{C}_i(m) \tag{32}$$

for $m = 0, 1, 2, \dots, M_2 - 1$ **do**

// Approximate $\mathbb{Q}^i_{\mathbf{K},\kappa}(\mathbf{x}_{\mathcal{N}_i^\kappa}(m), \mathbf{u}_{\mathcal{N}_i^\kappa}(m))$

Starting from $\bar{\mathbf{x}}^m(1)$ with control policy $\mathbf{K}(t)$ in a subroutine, resulting to a sample sequence: $\{\bar{\mathbf{x}}^m(1), \bar{\mathbf{u}}^m(1)), (\bar{\mathbf{x}}^m(2), \bar{\mathbf{u}}^m(2))), \dots, (\bar{\mathbf{x}}^m(l_2), \bar{\mathbf{u}}^m(l_2))\}$. For each agent i,

$$\widehat{\mathbb{Q}}^i_{\mathbf{K},\kappa}(\mathbf{x}_{\mathcal{N}_i^\kappa}(m), \mathbf{u}_{\mathcal{N}_i^\kappa}(m)) = \sum_{\tau=1}^{l_2} (\bar{c}_i(\tau) - \widehat{C}_i) = \sum_{\tau=1}^{l_2} [\bar{\mathbf{x}}_{\mathcal{N}_i}^m(\tau)^{\mathrm{T}} \widetilde{Q}_i \bar{\mathbf{x}}_{\mathcal{N}_i}^m(\tau) + \bar{\mathbf{u}}_i^m(\tau)^{\mathrm{T}} \widetilde{R}_i \bar{\mathbf{u}}_i^m(\tau) - \widehat{C}_i] \tag{33}$$

end for

// Policy Improvement

Calculate approximated gradient

$$\widehat{h}_i(t) = \frac{1}{M_2} \sum_{m=0}^{M_2-1} \nabla_{K_i(t)} \pi_{K_i(t)}(u_i(m) | x_{\mathcal{N}_i^r}(m)) \frac{1}{n} \sum_{j \in N_i^\kappa} \widehat{\mathbb{Q}}^j_{\mathbf{K},\kappa}(x_{N_j^\kappa}(m), u_{N_j^\kappa}(m)) \tag{34}$$

Conduct gradient step $\mathbf{K}_i(t+1) = \mathbf{K}_i(t) - \eta \widehat{h}_i(t)$.

References

[1] R. Zhang, W. Li, and N. Li, "On the optimal control of network lqr with spatially-exponential decaying structure," *arXiv preprint arXiv:2209.14376*, 2022.

[2] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 2012.

[3] J. Bu, A. Mesbahi, M. Fazel, and M. Mesbahi, "Lqr through the lens of first order methods: Discrete-time case," *arXiv preprint arXiv:1907.08921*, 2019.

[4] A. Beck, *First-order methods in optimization*. SIAM, 2017.

[5] Z. Yang, Y. Chen, M. Hong, and Z. Wang, "Provably global convergence of actor-critic: A case for linear quadratic regulator with ergodic cost," *Proc. Advances in Neural Inf. Process. Syst.*, vol. 32, 2019.

[6] M. Fazel, R. Ge, S. Kakade, and M. Mesbahi, "Global convergence of policy gradient methods for the linear quadratic regulator," in *Proc. Int. Conf. Mach. Learn.* PMLR, 2018, pp. 1467–1476.

[7] G. Qu, A. Wierman, and N. Li, "Scalable reinforcement learning of localized policies for multi-agent networked systems," in *Learn. Dyna. Cont.* PMLR, 2020, pp. 256–266.