Supplementary material: Exploiting inter-agent coupling information for efficient model-free reinforcement learning of cooperative LQR

Appendix A. Proof of Lemma 3.1

Proof needs to be written in $u(t), x(t)$. We prove a stronger version of the lemma that holds irrespective of the linear dynamics and quadratic cost assumption. For some $i, j \in \mathcal{V}$, let $j \in \mathcal{I}_Q^i$. For the sake of contradiction, assume that $\exists$ a $k \in \mathcal{R}_{SO}^j$ such that $k \notin \mathcal{I}_Q^i$. By the definition of $\mathcal{I}_Q^i$, $j \in \mathcal{I}_Q^i$ implies that for some some $t' \geq t$, $\exists$ a function (or composition of functions) $f : \mathcal{S} \times \mathcal{U} \to \mathbb{R}$ such that

$$c_i(x_{\mathcal{I}_C^i}(t'), u_{\mathcal{I}_C^i}(t')) = f(x_j(t), u_j(t), \bigcup_{g \in \mathcal{I}_Q^i \setminus j} \{x_g(\cdot), u_g(\cdot)\}). \tag{15}$$

Recall that the control $u_j(t) \in \mathcal{U}$ depends only on its partial observation $o_j(t)$, current state $x_j(t)$, and local policy $\pi_j(\cdot)$. Therefore, $\exists$ a function $g_j : \mathcal{Z}_j \to P(\mathcal{U}_j)$ such that

$$u_j(t) \sim g_j(o_j(t)) = g_j(\{x_m(t)\}_{m \in \mathcal{I}_O^j}) \tag{16}$$

Similarly, due to the Markovian assumption for each $x_j(t)$, $\exists$ a mapping $h_j : \prod_{n \in \mathcal{I}_S^j} \mathcal{S}_n \times \prod_{n \in \mathcal{I}_S^j} \mathcal{U}_n \to P(\mathcal{S}_j)$ such that

$$x_j(t) \sim h_j(\{x_n(t-1)\}_{n \in \mathcal{I}_S^j}, \{u_n(t-1)\}_{n \in \mathcal{I}_S^j}) \tag{17}$$

Using (16) and (17), (15) can be rewritten as

$$c_i(x_{\mathcal{I}_C^i}(t'), u_{\mathcal{I}_C^i}(t')) = f(x_j(t), u_j(t), \bigcup_{g \in \mathcal{I}_Q^i \setminus j} x_g, u_g) \tag{18}$$

$$= f(h_j(\{x_n(t-1)\}_{n \in \mathcal{I}_S^j}, \{u_n(t-1)\}_{n \in \mathcal{I}_S^j}), g_j(\{x_m(t)\}_{m \in \mathcal{I}_O^j}), \bigcup_{g \in \mathcal{I}_Q^i \setminus j} \{x_g(\cdot), u_g(\cdot)\}) \tag{19}$$

$$= f(h_j(\{x_n(t-1), u_n(t-1)\}_{n \in \mathcal{I}_S^j}), g_j(\{\{x_l(t-1), u_l(t-1)\}_{l \in \mathcal{I}_S^m}\}_{m \in \mathcal{I}_O^j}), \bigcup_{g \in \mathcal{I}_Q^i \setminus j} \{x_g(\cdot), u_g(\cdot)\}) \tag{20}$$

On recursive expansion of (20), it is straightforward to verify that $c_i(x_{\mathcal{I}_C^i}(t'), u_{\mathcal{I}_C^i}(t'))$ depends on $\{x_s(t''), u_s(t'')\}_{s \in \mathcal{R}_{SO}^j}$, for some $t'' \leq t \leq t'$. Thus, $i \in \mathcal{I}_{\text{GD}}^s \ \forall\ s \in \mathcal{R}_{SO}^j$ which implies that $s \in \mathcal{I}_Q^i \ \forall\ s \in \mathcal{R}_{SO}^j$. But as $k \in \mathcal{R}_{SO}^j$, $k \in \mathcal{I}_Q^i$ which is a contradiction. Therefore, our assumption is false and hence if $j \in \mathcal{I}_Q^i$, then $\forall\ k \in \mathcal{R}_{SO}^j$, $k \in \mathcal{I}_Q^i$ as required. ■

Appendix B. Proof of Theorem 3.1

Proof For the networked system, observe that the individual cost-to-go for each agent Q_i is dependent on the global state and control due to the long-term inter-agent dependencies between the agents. Recall that

$$Q_i(x,u) = c_i(x_{\mathcal{I}_C^i}, u_{\mathcal{I}_C^i}) + \mathbb{E}\left[\sum_{t=1}^{T} c_i(x_{\mathcal{I}_C^i}(t), u_{\mathcal{I}_C^i}(t))\right]. \tag{21}$$

For LTI dynamics (1) and quadratic cost (2), (21) can be rewritten as

$$\begin{aligned}
Q_i(x,u) &= \begin{bmatrix} x_{\mathcal{I}_C^i}(t) \\ u_{\mathcal{I}_C^i}(t) \end{bmatrix}^\intercal \begin{bmatrix} S_i & 0 \\ 0 & R_i \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}_C^i}(t) \\ u_{\mathcal{I}_C^i}(t) \end{bmatrix} + \mathbb{E}_{w(t),\eta(t)}\left[\begin{bmatrix} x_{\mathcal{I}_C^i}(t+1) \\ u_{\mathcal{I}_C^i}(t+1) \end{bmatrix}^\intercal \begin{bmatrix} S_i & 0 \\ 0 & R_i \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}_C^i}(t+1) \\ u_{\mathcal{I}_C^i}(t+1) \end{bmatrix} \right. \\
&\left. + \mathbb{E}_{w(t+1),\eta(t+1)} \begin{bmatrix} x_{\mathcal{I}_C^i}(t+2) \\ u_{\mathcal{I}_C^i}(t+2) \end{bmatrix}^\intercal \begin{bmatrix} S_i & 0 \\ 0 & R_i \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}_C^i}(t+2) \\ u_{\mathcal{I}_C^i}(t+2) \end{bmatrix} + \mathbb{E}\left[\cdots\right] \right] = \\
&\sum_{j,k\in\mathcal{I}_C^i} \left[(x_j(t))^\intercal S_{jk}(x_k(t)) + (u_j(t))^\intercal R_{jk}(u_k(t)) + \left[\sigma_w^2 \mathrm{Tr}\left(S_i\right) + \sigma_\eta^2 \mathrm{Tr}\left(R_i\right)\right]_{j\in\mathcal{I}_C^i} + \right. \\
&\left[x^\intercal_{\mathcal{I}_S^j}(t) A_j^\intercal S_i A_j x_{\mathcal{I}_S^j}(t) + u^\intercal_{\mathcal{I}_S^j}(t) B_j^\intercal S_i B_j u_{\mathcal{I}_S^j}(t) + 2 x^\intercal_{\mathcal{I}_S^j}(t) A_j^\intercal S_i B_j u_{\mathcal{I}_S^j}(t) + x^\intercal_{\mathcal{I}_O^j}(t) K_j^\intercal R_i K_j x_{\mathcal{I}_O^j}(t) \right]_{j\in\mathcal{I}_C^i} \\
&\left. + \sigma_\eta^2 \mathrm{Tr}\left(B_j^\intercal S_i B_j \mathbb{I}_{n_u|\mathcal{I}_S^j|}\right) + 2\mathrm{Tr}\left(A_j^\intercal S_i B_j w_k(t) \eta_l^\intercal(t)\right)_{\substack{k\in\mathcal{I}_S^j \\ l\in\mathcal{I}_O^j}} + + \sigma_w^2 \mathrm{Tr}\left(A_j^\intercal S_i A_j \mathbb{I}_{n_x|\mathcal{I}_S^j|}\right) + \cdots \right]
\end{aligned} \tag{22}$$

Therefore, from (22), it is clear that for time-invariant inter-agent couplings, the $Q_i(\cdot)$ for each $i \in \mathcal{V}$ depends on its neighbors in the cost graph which in turn depend on their neighbors in the state, and observation graphs and so on. In other words, $\forall\ i \in \mathcal{V}$, $Q_i(\cdot)$ depends on a subset of agents $\mathcal{I}_Q^i := \{\mathcal{I}_C^i \cup \{\mathcal{R}_{SO}^k\}_{k\in\mathcal{I}_C^i}\} = \{\mathcal{R}_{SO}^k\}_{k\in\mathcal{I}_C^i}$. By Lemma 3.1, we have that $\mathcal{I}_Q^i$ is closed under $\mathcal{R}_S O$ which implies that the information of agents in $\mathcal{I}_Q^i$ is sufficient to exactly compute the the future costs of agent i. Therefore, it follows that $Q_i(x(t), u(t)) = Q_i(x_{\mathcal{I}_Q^i}(t), u_{\mathcal{I}_Q^i}(t))$ as required.
■

Appendix C. Proof of Theorem 3.2

Proof Recall that

$$
\begin{aligned}
Q^\pi(x,u) &= \mathbb{E}_\pi\left[\sum_{i=1}^{N}\sum_{t=0}^{\infty} c_i(x_{\mathcal{I}_C^i}(t), u_{\mathcal{I}_C^i}(t))|x(0)=x, u(0)=u\right] \\
&= \mathbb{E}_\pi\Bigg[\sum_{j\in\mathcal{I}_{\text{GD}}^i}\sum_{t=0}^{\infty} c_j(x_{\mathcal{I}_C^j}(t), u_{\mathcal{I}_C^j}(t))|x(0)=x, u(0)=u\Bigg] \\
&\qquad + \mathbb{E}_\pi\Bigg[\sum_{j\backslash\mathcal{I}_{\text{GD}}^i}\sum_{t=0}^{\infty} c_j(x_{\mathcal{I}_C^j}(t), u_{\mathcal{I}_C^j}(t))|x(0)=x, u(0)=u\Bigg] \\
&= \sum_{j\in\mathcal{I}_{\text{GD}}^i} Q_j^\pi(x_{\mathcal{I}_Q^j}, u_{\mathcal{I}_Q^j}) + \sum_{k\backslash\mathcal{I}_{\text{GD}}^i} Q_k^\pi(x_{\mathcal{I}_Q^k}, u_{\mathcal{I}_Q^k}) = \widehat{Q}_i^\pi(x_{\mathcal{I}_{\widehat{Q}}^j}, u_{\mathcal{I}_{\widehat{Q}}^j}) + \bar{Q}_i^\pi(x\mathcal{I}_{\bar{Q}}^i, u_{\mathcal{I}^i\bar{Q}}),
\end{aligned}
\tag{23}
$$

where $\bar{Q}_i^\pi(x\mathcal{I}_{\bar{Q}}^i, u_{\mathcal{I}^i\bar{Q}}) = Q^\pi(x,u) - \widehat{Q}_i^\pi(x_{\mathcal{I}_{\widehat{Q}}^j}, u_{\mathcal{I}_{\widehat{Q}}^j}) = \sum_{k\backslash\mathcal{I}_{\text{GD}}^i} Q_k^\pi(x_{\mathcal{I}_Q^k}, u_{\mathcal{I}_Q^k})$. From Theorem 3.1, the reward of each agent $i \in \mathcal{V}$ depends on $x_j(t)$, $u_j(t)\ \forall\ j \in \mathcal{I}_Q^i$ and $\mathcal{E}_{\text{GD}} = \mathcal{E}_{\text{VD}}^\intercal$ by definition of $\mathcal{G}_{\text{GD}}$. Therefore, if $j \notin \mathcal{I}_{\text{GD}}^i$, then $i \notin \mathcal{I}_Q^j$. Hence, $\sum_{j\backslash\mathcal{I}_{\text{GD}}^i} c_j(x_{\mathcal{I}_C^j}(t), u_{\mathcal{I}_C^j}(t))$ is independent of $u_i(t)$ and thus K_i. It then follows that $Q_j^\pi(\cdot)$, is independent of K_i, $\forall\ j \notin \mathcal{I}_{\text{GD}}^i$, which implies

$$
\begin{aligned}
\nabla_{K_i}\bar{Q}_i^\pi &= \nabla_{K_i}\mathbb{E}_\pi\Bigg[\sum_{j\backslash\mathcal{I}_{\text{GD}}^i}\sum_{t=0}^{\infty} c_j(x_{\mathcal{I}_C^j}(t), u_{\mathcal{I}_C^j}(t))|x(0)=x, u(0)=u\Bigg] \\
&\overset{(a)}{=} \mathbb{E}_\pi\left[\nabla_{K_i}\sum_{j\backslash\mathcal{I}_{\text{GD}}^i}\sum_{t=0}^{\infty} c_j(x_{\mathcal{I}_C^j}(t), u_{\mathcal{I}_C^j}(t))|x(0)=x, u(0)=u\right] = 0,
\end{aligned}
\tag{24}
$$

where (a) in (24) is obtained by interchanging the derivative and integral assuming that each $Q_j^\pi(\cdot)$ is sufficiently smooth in state and control. Hence, the gradient of the global action value function with respect to K_i is given by $\nabla_{K_i}Q^\pi(s,a) = \nabla_{K_i}[\widehat{Q}_i^\pi + \bar{Q}_i^\pi] = \nabla_{K_i}\hat{Q}_i^\pi,$ as required. ■

Appendix D. Proof of Proposition 4.1

Proof From (9), we have

$$
\widehat{Q}_i^\pi(x_{\mathcal{I}_{\widehat{Q}}^i}, u_{\mathcal{I}_{\widehat{Q}}^i}) = \begin{bmatrix} x_{\mathcal{I}_{\widehat{Q}}^i}(t) \\ u_{\mathcal{I}_{\widehat{Q}}^i}(t) \end{bmatrix} \begin{bmatrix} S_{\mathcal{I}_{\widehat{Q}}^i} & 0 \\ 0 & R_{\mathcal{I}_{\widehat{Q}}^i} \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}_{\widehat{Q}}^i}(t) \\ u_{\mathcal{I}_{\widehat{Q}}^i}(t) \end{bmatrix} + \mathbb{E}\left[\widehat{Q}_i(x_{\mathcal{I}_{\widehat{Q}}^i}(t+1), u_{\mathcal{I}_{\widehat{Q}}^i}(t+1))\right]. \tag{25}
$$

Then, the expected future Q-value can be rewritten as

$$
\begin{aligned}
&\mathbb{E}\left[\widehat{Q}_i(x_{\mathcal{I}^i_{\widehat{Q}}}(t+1), u_{\mathcal{I}^i_{\widehat{Q}}}(t+1))\right] \\
&= \mathbb{E}\left[\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t+1) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t+1) \end{bmatrix} \begin{bmatrix} S_{\mathcal{I}^i_{\widehat{Q}}} & 0 \\ 0 & R_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t+1) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t+1) \end{bmatrix}\right] + \mathbb{E}\left[\mathbb{E}\left[\widehat{Q}_i(x_{\mathcal{I}^i_{\widehat{Q}}}(t+2), u_{\mathcal{I}^i_{\widehat{Q}}}(t+2))\right]\right] \\
&= \mathbb{E}\left[(A_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t) + B_{\mathcal{I}^i_{\widehat{Q}}} u_{\mathcal{I}^i_{\widehat{Q}}}(t) + w_{\mathcal{I}^i_{\widehat{Q}}}(t))^\intercal S_{\mathcal{I}^i_{\widehat{Q}}} (A_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t) + B_{\mathcal{I}^i_{\widehat{Q}}} u_{\mathcal{I}^i_{\widehat{Q}}}(t) + w_{\mathcal{I}^i_{\widehat{Q}}}(t))\right] + \\
&\mathbb{E}\Big[(K_{\mathcal{I}^i_{\widehat{Q}}}(A_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t) + B_{\mathcal{I}^i_{\widehat{Q}}} u_{\mathcal{I}^i_{\widehat{Q}}}(t) + w_{\mathcal{I}^i_{\widehat{Q}}}(t)))^\intercal R_{\mathcal{I}^i_{\widehat{Q}}} (K_{\mathcal{I}^i_{\widehat{Q}}}(A_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t) + B_{\mathcal{I}^i_{\widehat{Q}}} u_{\mathcal{I}^i_{\widehat{Q}}}(t) + w_{\mathcal{I}^i_{\widehat{Q}}}(t)))\Big] \\
&+ \mathbb{E}\left[\mathbb{E}\left[\widehat{Q}_i(x_{\mathcal{I}^i_{\widehat{Q}}}(t+2), u_{\mathcal{I}^i_{\widehat{Q}}}(t+2))\right]\right] \\
&= (A_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t) + B_{\mathcal{I}^i_{\widehat{Q}}} u_{\mathcal{I}^i_{\widehat{Q}}}(t))^\intercal S_{\mathcal{I}^i_{\widehat{Q}}} (A_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t) + B_{\mathcal{I}^i_{\widehat{Q}}} u_{\mathcal{I}^i_{\widehat{Q}}}(t)) + \sigma_w^2 \mathrm{Tr}\left(S_{\mathcal{I}^i_{\widehat{Q}}} + K^\intercal_{\mathcal{I}^i_{\widehat{Q}}} R_{\mathcal{I}^i_{\widehat{Q}}} K_{\mathcal{I}^i_{\widehat{Q}}}\right) \\
&+ (K_{\mathcal{I}^i_{\widehat{Q}}}(A_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t) + B_{\mathcal{I}^i_{\widehat{Q}}} u_{\mathcal{I}^i_{\widehat{Q}}}(t)))^\intercal R_{\mathcal{I}^i_{\widehat{Q}}} (K_{\mathcal{I}^i_{\widehat{Q}}}(A_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t) + B_{\mathcal{I}^i_{\widehat{Q}}} u_{\mathcal{I}^i_{\widehat{Q}}}(t))) \\
&+ \mathbb{E}\left[\mathbb{E}\left[\widehat{Q}_i(x_{\mathcal{I}^i_{\widehat{Q}}}(t+2), u_{\mathcal{I}^i_{\widehat{Q}}}(t+2))\right]\right]
\end{aligned} \tag{26}
$$

$$
\begin{aligned}
&= \begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t) \end{bmatrix} \begin{bmatrix} A^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \\ B^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} (S_{\mathcal{I}^i_{\widehat{Q}}} + K^\intercal_{\mathcal{I}^i_{\widehat{Q}}} R_{\mathcal{I}^i_{\widehat{Q}}} K_{\mathcal{I}^i_{\widehat{Q}}}) \begin{bmatrix} A_{\mathcal{I}^i_{\widehat{Q}}} & B_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t) \end{bmatrix} + \sigma_w^2 \begin{bmatrix} \mathbb{I} \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal \begin{bmatrix} S_{\mathcal{I}^i_{\widehat{Q}}} & 0 \\ 0 & R_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \begin{bmatrix} \mathbb{I} \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \\
&+ \mathbb{E}\left[\mathbb{E}\left[\widehat{Q}_i(x_{\mathcal{I}^i_{\widehat{Q}}}(t+2), u_{\mathcal{I}^i_{\widehat{Q}}}(t+2))\right]\right].
\end{aligned} \tag{27}
$$

Recursive expansion of (27) yields

$$
\widehat{Q}_i(x_{\mathcal{I}^i_{\widehat{Q}}}, u_{\mathcal{I}^i_{\widehat{Q}}}) = \begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t) \end{bmatrix} \widehat{Q}_i \begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t) \end{bmatrix} + \sigma_w^2 \begin{bmatrix} \mathbb{I} \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal \widehat{Q}_i \begin{bmatrix} \mathbb{I} \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}, \tag{28}
$$

where with a slight abuse of notation how is the 2nd term obtained?

$$
\widehat{Q}_i = \begin{bmatrix} S_{\mathcal{I}^i_{\widehat{Q}}} & 0 \\ 0 & R_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} + \begin{bmatrix} A^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \\ B^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \mathcal{L}\left(A_{\mathcal{I}^i_{\widehat{Q}}} + B_{\mathcal{I}^i_{\widehat{Q}}} K_{\mathcal{I}^i_{\widehat{Q}}}, S_{\mathcal{I}^i_{\widehat{Q}}} + K^\intercal_{\mathcal{I}^i_{\widehat{Q}}} R_{\mathcal{I}^i_{\widehat{Q}}} K_{\mathcal{I}^i_{\widehat{Q}}}\right) \begin{bmatrix} A_{\mathcal{I}^i_{\widehat{Q}}} & B_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix},
$$

$\mathcal{L}(X, Y)$ is the analytical solution of the discrete time Lyapunov equation $\mathcal{P} = X\mathcal{P}X^\intercal + Y$. ■

Appendix E. Proof of Lemma 5.1

Proof

(a) **Necessary condition.** Assume that $\mathcal{I}^i_{\widehat{Q}} \subset \mathcal{V}$. This implies that $\exists\, k \in \mathcal{V}$ such that $k \notin \bigcup_{j\in\mathcal{I}^i_{\text{GD}}} \mathcal{I}^j_Q$ i.e., $k \notin \mathcal{I}^i_{\text{GD}}$, and $k \notin \mathcal{I}^j_Q, \forall\, j \in \mathcal{I}^i_{\text{GD}}$. By definition, $k \notin \mathcal{I}^i_{\text{GD}}$ if $i \notin \mathcal{I}^k_Q$ which implies $i \notin \mathcal{I}^k_C$, and $i \notin \{\mathcal{R}^l_{SO}\}_{l\in\mathcal{I}^k_C}$. Similarly, $k \notin \mathcal{I}^j_Q, \forall\, j \in \mathcal{I}^i_{\text{GD}}$ if $k \notin \mathcal{I}^j_C$, and $k \notin \{\mathcal{R}^m_{SO}\}_{m\in\mathcal{I}^j_C}$, $\forall\, j \in \mathcal{I}^i_{\text{GD}}$. However, if $j \in \mathcal{I}^i_{\text{GD}}$, implies $i \in \mathcal{I}^j_Q$ i.e., either $i \in \mathcal{I}^j_C$ or $i \in \{\mathcal{R}^m_{SO}\}_{m\in\mathcal{I}^j_C}$. Since, $\forall\, j \in \mathcal{I}^i_{\text{GD}}$, $i \in \mathcal{I}^j_C$, and $k \notin \mathcal{I}^j_C$ implies that $\mathcal{I}^i_{C^\intercal} \cap \mathcal{I}^k_{C^\intercal} = \emptyset$. Also, as $i \notin \{\mathcal{R}^l_{SO}\}_{l\in\mathcal{I}^k_C}$, and $k \notin \{\mathcal{R}^m_{SO}\}_{m\in\mathcal{I}^j_C}$ implies that $\mathcal{R}^i_{(SO)^\intercal} \cap \mathcal{R}^k_{(SO)^\intercal} = \emptyset$ as required.

Sufficient condition. We prove this using proof by contraposition. Assume that $\forall\, k \in \mathcal{V}$, $\mathcal{I}^i_{C^\intercal} \cap \mathcal{I}^k_{C^\intercal} \neq \emptyset$, or $\mathcal{R}^i_{(SO)^\intercal} \cap \mathcal{R}^k_{(SO)^\intercal} \neq \emptyset$.

Case 1: If $\mathcal{I}^i_{C^\intercal} \cap \mathcal{I}^k_{C^\intercal} \neq \emptyset$, then $\exists\, j \in \mathcal{V}$ such that $\{i,k\} \in \mathcal{I}^j_C$ which implies $j \in \mathcal{I}^i_{\text{GD}}$ and $k \in \mathcal{I}^j_Q$. Therefore, $k \in \mathcal{I}^i_{\widehat{Q}}$ which implies $\mathcal{I}^i_{\widehat{Q}} = \mathcal{V}$.

Case 2: If $\mathcal{R}^i_{(SO)^\intercal} \cap \mathcal{R}^k_{(SO)^\intercal} \neq \emptyset$, then $\exists\, j \in \mathcal{V}$ such that $\{i,k\} \in \mathcal{R}^j_{SO}$. Since, by definition, $j \in \mathcal{I}^j_C$ implies that $\{i,k\} \in \mathcal{I}^j_Q$ and $j \in \mathcal{I}^i_{\text{GD}} \cap \mathcal{I}^k_{\text{GD}}$. Hence, as $j \in \mathcal{I}^i_{\text{GD}}$, and $k \in \mathcal{I}^j_Q$, we have that $k \in \mathcal{I}^i_{\widehat{Q}}$ which implies that $\mathcal{I}^i_{\widehat{Q}} = \mathcal{V}$.

In either case, $\mathcal{I}^i_{\widehat{Q}} = \mathcal{V}$. Therefore, $\forall\, k \in \mathcal{V}$, if $\mathcal{I}^i_{C^\intercal} \cap \mathcal{I}^k_{C^\intercal} \neq \emptyset$, or $\mathcal{R}^i_{(SO)^\intercal} \cap \mathcal{R}^k_{(SO)^\intercal} \neq \emptyset$, then $\mathcal{I}^i_{\widehat{Q}} = \mathcal{V}$. The sufficient condition follows by contraposition as required.

(b) **Necessary condition.** Consider an $i \in \mathcal{V}$ and assume that $\exists$ an $j \in \mathcal{I}^i_{\text{GD}}$, such that $\mathcal{I}^j_Q \subset \mathcal{I}^i_{\widehat{Q}}$. This implies that $\exists\, k \in \mathcal{I}^i_{\widehat{Q}}$ such that $k \notin \mathcal{I}^j_Q$, and $k \in \bigcup_{h\in\mathcal{I}^i_{\text{GD}}\setminus\{j\}} \mathcal{I}^h_Q$. If $k \notin \mathcal{I}^j_Q$, then by definition, $k \notin \mathcal{I}^j_C$, and $k \notin \{\mathcal{R}^l_{SO}\}_{l\in\mathcal{I}^j_C}$. But, as $k \in \bigcup_{h\in\mathcal{I}^i_{\text{GD}}\setminus\{j\}} \mathcal{I}^h_Q$, implies that $\exists\, h \in \mathcal{I}^i_{\text{GD}} \setminus \{j\}$ such that either $k \in \mathcal{I}^h_C$ or $k \in \{\mathcal{R}^m_{SO}\}_{m\in\mathcal{I}^h_C}$.

Case 1 Let $k \in \mathcal{I}^h_C$. Then, as $h \in \mathcal{I}^i_{\text{GD}}$, either $i \in \mathcal{I}^h_C$, or $i \in \{\mathcal{R}^l_{SO}\}_{l\in\mathcal{I}^h_C}$.

- If $i \in \mathcal{I}^h_C$, then $\mathcal{I}^i_{C^\intercal} \cap \mathcal{I}^k_{C^\intercal} = \{h\} \neq \emptyset$. or,
- If $i \in \{\mathcal{R}^l_{SO}\}_{l\in\mathcal{I}^j_C}$, then $\exists$ an $m \in \mathcal{I}^h_C \cap \mathcal{R}^i_{(SO)^\intercal}$. Hence, $\mathcal{I}^m_{C^\intercal} \cap \mathcal{I}^k_{C^\intercal} = \{h\} \neq \emptyset$.

Case 2 Let $k \in \{\mathcal{R}^m_{SO}\}_{m\in\mathcal{I}^h_C}$. Then, $\exists$ an $p \in \mathcal{I}^h_C \cap \mathcal{R}^k_{(SO)^\intercal}$, and as $h \in \mathcal{I}^i_{\text{GD}}$, either $i \in \mathcal{I}^h_C$, or $i \in \{\mathcal{R}^l_{SO}\}_{l\in\mathcal{I}^h_C}$.

- If $i \in \mathcal{I}^j_C$, then $\mathcal{I}^i_{C^\intercal} \cap \mathcal{I}^p_{C^\intercal} = \{h\} \neq \emptyset$. or,
- If $i \in \{\mathcal{R}^l_{SO}\}_{l\in\mathcal{I}^h_C}$, then $\exists$ an $m \in \mathcal{I}^h_C \cap \mathcal{R}^i_{(SO)^\intercal}$. Hence, $\mathcal{I}^m_{C^\intercal} \cap \mathcal{I}^p_{C^\intercal} = \{h\} \neq \emptyset$.

Therefore, in either case we conclude that if $\mathcal{I}^i_Q \subset \mathcal{I}^i_{\widehat{Q}}$, then $p \in \mathcal{R}^k_{(SO)^\intercal}$, $m \in \mathcal{R}^i_{(SO)^\intercal}$, such that $\mathcal{I}^m_{C^\intercal} \cap \mathcal{I}^p_{C^\intercal} \subset \mathcal{I}^i_{\text{GD}}$.

Sufficient condition. Consider an $i \in \mathcal{V}$ and assume that $\exists$ an $j \in \mathcal{I}^i_{\text{GD}}$ for which $\exists$ a $k \in \mathcal{V}\setminus\mathcal{I}^j_Q$. Let $\exists\, h \in \mathcal{I}^i_{\text{GD}}$, $m \in \mathcal{R}^i_{(SO)^\intercal}$, and $p \in \mathcal{R}^k_{(SO)^\intercal}$, such that $h \in \mathcal{I}^m_{C^\intercal} \cap \mathcal{I}^p_{C^\intercal}$. Hence, as $p \in \mathcal{I}^h_C$, by definition $k \in \mathcal{I}^h_Q$. But, as $h \in \mathcal{I}^i_{\text{GD}}$, we have that $k \in \mathcal{I}^i_{\widehat{Q}}$. However, $k \notin \mathcal{I}^j_Q$ which implies that $k \in \mathcal{I}^i_{\widehat{Q}} \setminus \mathcal{I}^i_Q$ or $\mathcal{I}^i_Q \subset \mathcal{I}^i_{\widehat{Q}}$ as required.

■

Appendix F. Proof of Theorem 5.1

Employing a *linear architecture*, (**??**) can be expressed as

$$\sum_{j\in\mathcal{I}^i_{\mathrm{GD}}} r_j(t) =$$
$$\lambda + \left[\operatorname{svec}\left(\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t)\end{bmatrix}\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t)\end{bmatrix}^\intercal\right) - \mathbb{E}\left[\operatorname{svec}\left(\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t+1) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t+1)\end{bmatrix}\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t+1) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t+1)\end{bmatrix}^\intercal\right)\right]\right]^\intercal \operatorname{svec}(P_{\mathcal{I}^i_{\widehat{Q}}}), \tag{29}$$

where $\lambda \in \mathbb{R}$ is a free parameter to satisfy the fixed point equation. Let $\lambda = \left\langle P_i, \begin{bmatrix} \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} & \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} K^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \\ K_{\mathcal{I}^i_{\widehat{Q}}}\Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} & K_{\mathcal{I}^i_{\widehat{Q}}}\Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} K^\intercal_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}\right\rangle$.

For brevity, in the remainder of the proof denote $\phi_t = \operatorname{svec}\left(\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t)\end{bmatrix}\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t)\end{bmatrix}^\intercal\right)$, $\psi_t = \operatorname{svec}\left(\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ K_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t)\end{bmatrix}\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ K_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t)\end{bmatrix}^\intercal\right)$, $f = \operatorname{svec}\left(\begin{bmatrix} \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} & \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} K^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \\ K_{\mathcal{I}^i_{\widehat{Q}}}\Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} & K_{\mathcal{I}^i_{\widehat{Q}}}\Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} K^\intercal_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}\right)$, and $\xi_t = \mathbb{E}\left[\operatorname{svec}\left(\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t+1) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t+1)\end{bmatrix}\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t+1) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t+1)\end{bmatrix}^\intercal\right)\right]$.

Assuming that we have a single trajectory of $\left\{x_{\mathcal{I}^i_{\widehat{Q}}}(t), u_{\mathcal{I}^i_{\widehat{Q}}}(t), x_{\mathcal{I}^i_{\widehat{Q}}}(t+1)\right\}_{t=1}^T$, (98) can be expressed in matrix form as

$$\mathbf{r} = (\mathbf{\Phi} - \mathbf{\Xi} + \mathbf{F})p, \tag{30}$$

where

$$\mathbf{\Phi} = \begin{bmatrix}\phi_1^\intercal \\ \phi_2^\intercal \\ \vdots \\ \phi_T^\intercal\end{bmatrix}, \ \mathbf{\Xi} = \begin{bmatrix}\xi_1^\intercal \\ \xi_2^\intercal \\ \vdots \\ \xi_T^\intercal\end{bmatrix}, \ \mathbf{r} = \begin{bmatrix} r(1) \\ r(2) \\ \vdots \\ r(T)\end{bmatrix}, \ \mathbf{F} = \begin{bmatrix} f^\intercal \\ f^\intercal \\ \vdots \\ f^\intercal\end{bmatrix}$$

Let

$$\psi_+ = \begin{bmatrix}\psi_2^\intercal \\ \psi_3^\intercal \\ \vdots \\ \psi_{T+1}^\intercal\end{bmatrix}.$$

Note that (30) is an instance of the error-in-variables least square problem whose solution is given by

$$\hat{p} = (\mathbf{\Phi}^\intercal(\mathbf{\Phi} - \mathbf{\Psi}_+ + \mathbf{F}))^{-1}\mathbf{\Phi}^\intercal\mathbf{r} \tag{31}$$

Rearranging the terms in (31) to obtain

$$\begin{aligned}
&\mathbf{\Phi}^\intercal(\mathbf{\Phi} - \mathbf{\Psi}_+ + \mathbf{F})\hat{p} = \mathbf{\Phi}^\intercal \mathbf{r} \\
&\Rightarrow \mathbf{\Phi}\hat{p} = \mathbf{\Phi}(\mathbf{\Phi}^\intercal\mathbf{\Phi})^{-1}\mathbf{\Phi}^\intercal(\mathbf{r} + (\mathbf{\Psi}_+ - \mathbf{F})\hat{p})
\end{aligned} \tag{32}$$

Define $P_{\mathbf{\Phi}} = \mathbf{\Phi}(\mathbf{\Phi}^\intercal\mathbf{\Phi})^{-1}\mathbf{\Phi}^\intercal$ as the orthogonal projection onto the columns of $\mathbf{\Phi}$. Combining (30), (32), and using the fact that $P_{\mathbf{\Phi}}\mathbf{\Phi} = \mathbf{\Phi}$ yields

$$P_{\mathbf{\Phi}}(\mathbf{\Phi} - \mathbf{\Xi} + \mathbf{F})(p - \hat{p}) = P_{\phi}(\mathbf{\Xi} - \mathbf{\Psi}_+)\hat{p} \tag{33}$$

Consider the i^{th} row of $\mathbf{\Phi} - \mathbf{\Xi} + \mathbf{F}$,

$$\begin{aligned}
&\text{svec}\Bigg(\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t) \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t) \end{bmatrix}^\intercal - \mathbb{E}\left[\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t+1) \\ K_{\mathcal{I}^i_{\widehat{Q}}}(x_{\mathcal{I}^i_{\widehat{Q}}}(t+1)) \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t+1) \\ K_{\mathcal{I}^i_{\widehat{Q}}}(x_{\mathcal{I}^i_{\widehat{Q}}}(t+1)) \end{bmatrix}^\intercal \right] \\
&\qquad\qquad + \begin{bmatrix} \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} & \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} K^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \\ K_{\mathcal{I}^i_{\widehat{\widehat{Q}}}} \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} & K_{\mathcal{I}^i_{\widehat{Q}}} \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} K^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \Bigg), \\
&\text{where } x_{\mathcal{I}^i_{\widehat{Q}}}(t+1) = A_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{Q}}(t) + B_{\mathcal{I}^i_{\widehat{Q}}} u_{\mathcal{I}^i_{\widehat{Q}}}(t) + \eta_{\mathcal{I}^i_{\widehat{Q}}}. \\
&= \text{svec}\Bigg(\begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t) \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t) \end{bmatrix}^\intercal - L \begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t) \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}^i_{\widehat{Q}}}(t) \\ u_{\mathcal{I}^i_{\widehat{Q}}}(t) \end{bmatrix}^\intercal L^\intercal - \begin{bmatrix} \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} & \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} K^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \\ K_{\mathcal{I}^i_{\widehat{Q}}} \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} & K_{\mathcal{I}^i_{\widehat{Q}}} \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} K^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \\
&\qquad\qquad + \begin{bmatrix} \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} & \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} K^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \\ K_{\mathcal{I}^i_{\widehat{Q}}} \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} & K_{\mathcal{I}^i_{\widehat{Q}}} \Sigma^x_{\mathcal{I}^i_{\widehat{Q}}} K^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \Bigg), \\
&\text{where } L = \begin{bmatrix} A_{\mathcal{I}^i_{\widehat{Q}}} & B_{\mathcal{I}^i_{\widehat{Q}}} \\ K_{\mathcal{I}^i_{\widehat{Q}}} A_{\mathcal{I}^i_{\widehat{Q}}} & K_{\mathcal{I}^i_{\widehat{Q}}} B_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \\
&= (\mathbb{I} - L \otimes_s L)\phi_t.
\end{aligned} \tag{34}$$

Combining (34) and (33) and assuming that $\mathbf{\Phi}$ is full column rank, we obtain

$$\begin{aligned}
&\mathbf{\Phi}(\mathbb{I} - L \otimes_s L)^\intercal(p - \hat{p}) = P_{\phi}(\mathbf{\Xi} - \mathbf{\Psi}_+)\hat{p} \\
&\Rightarrow (\mathbb{I} - L \otimes_s L)^\intercal(p - \hat{p}) = (\mathbf{\Phi}^\intercal\mathbf{\Phi})^{-1}\mathbf{\Phi}^\intercal(\mathbf{\Xi} - \mathbf{\Psi}_+)\hat{p}.
\end{aligned} \tag{35}$$

Let $\sigma_{\min}(\cdot)$ denote the minimum singular value of a matrix. Then, we have that

$$||(\mathbb{I} - L \otimes_s L)^\mathsf{T}(p - \hat{p})|| \geq \sigma_{\min}(\mathbb{I} - L \otimes_s L)||p - \hat{p}||, \tag{36}$$

$$\begin{aligned}
||(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-1}\mathbf{\Phi}^\mathsf{T}(\mathbf{\Xi} - \mathbf{\Psi}_+)\hat{p}|| &\leq \sigma_{\max}((\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}})||(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\mathsf{T}(\mathbf{\Xi} - \mathbf{\Psi}_+)\hat{p}|| \\
&= \lambda_{\max}((\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}})||(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\mathsf{T}(\mathbf{\Xi} - \mathbf{\Psi}_+)\hat{p}|| \\
&(\because \mathbf{\Phi}^\mathsf{T}\mathbf{\Phi} \text{ is symmetric and P.S.D,} (\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}} \text{ is symmetric and P.S.D.}) \\
&= \frac{||(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\mathsf{T}(\mathbf{\Xi} - \mathbf{\Psi}_+)\hat{p}||}{\lambda_{\min}((\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{\frac{1}{2}})} \\
&= \frac{||(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\mathsf{T}(\mathbf{\Xi} - \mathbf{\Psi}_+)\hat{p}||}{\sqrt{\lambda_{\min}(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})}} \\
&= \frac{||(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\mathsf{T}(\mathbf{\Xi} - \mathbf{\Psi}_+)\hat{p}||}{\sigma_{\min}(\mathbf{\Phi})}
\end{aligned} \tag{37}$$

Combining (39), (36), (37) yields

$$\begin{aligned}
&\sigma_{\min}(\mathbb{I} - L \otimes_s L)||p - \hat{p}|| \leq \frac{||(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\mathsf{T}(\mathbf{\Xi} - \mathbf{\Psi}_+)\hat{p}||}{\sigma_{\min}(\mathbf{\Phi})} \\
&\Rightarrow ||p - \hat{p}|| \leq \frac{||(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\mathsf{T}(\mathbf{\Xi} - \mathbf{\Psi}_+)\hat{p}||}{\sigma_{\min}(\mathbf{\Phi})\sigma_{\min}(\mathbb{I} - L \otimes_s L)} \\
&\leq \frac{||(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\mathsf{T}(\mathbf{\Xi} - \mathbf{\Psi}_+)||||p - \hat{p}||}{\sigma_{\min}(\mathbf{\Phi})\sigma_{\min}(\mathbb{I} - L \otimes_s L)} + \frac{||(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\mathsf{T}(\mathbf{\Xi} - \mathbf{\Psi}_+)p||}{\sigma_{\min}(\mathbf{\Phi})\sigma_{\min}(\mathbb{I} - L \otimes_s L)} \\
&\qquad \text{(By triangle inequality and Cauchy-Scwartz inequality)}
\end{aligned} \tag{38}$$

If $\frac{||(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\mathsf{T}(\mathbf{\Xi} - \mathbf{\Psi}_+)||}{\sigma_{\min}(\mathbf{\Phi})\sigma_{\min}(\mathbb{I} - L \otimes_s L)} < \frac{1}{2}$, then

$$||p - \hat{p}|| \leq 2\frac{||(\mathbf{\Phi}^\mathsf{T}\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\mathsf{T}(\mathbf{\Xi} - \mathbf{\Psi}_+)p||}{\sigma_{\min}(\mathbf{\Phi})\sigma_{\min}(\mathbb{I} - L \otimes_s L)}. \tag{39}$$

F.1. Lower bound on $\sigma_{\min}(\mathbf{\Phi})$

Let y be an arbitrary vector on a unit sphere $\mathcal{S}^{n_{\hat{x}}+n_{\hat{u}}-1}$. Define a process $Z_t = \langle \phi_t, y \rangle$, the filtration $\mathcal{F}_t = \sigma(u_{\mathcal{I}^i_{\hat{Q}}}(\tau), w_{\mathcal{I}^i_{\hat{Q}}}(\tau-1)\}_{\tau=1}^T$.

Then,

$$\begin{bmatrix} x_{\mathcal{I}^i_{\hat{Q}}}(t+1) \\ u_{\mathcal{I}^i_{\hat{Q}}}(t+1) \end{bmatrix} = \begin{bmatrix} A_{\mathcal{I}^i_{\hat{Q}}} x_{\mathcal{I}^i_{\hat{Q}}}(t) + B_{\mathcal{I}^i_{\hat{Q}}} u_{\mathcal{I}^i_{\hat{Q}}}(t) + w_{\mathcal{I}^i_{\hat{Q}}}(t) \\ K'_{\mathcal{I}^i_{\hat{Q}}}(A_{\mathcal{I}^i_{\hat{Q}}} x_{\mathcal{I}^i_{\hat{Q}}}(t) + B_{\mathcal{I}^i_{\hat{Q}}} u_{\mathcal{I}^i_{\hat{Q}}}(t) + w_{\mathcal{I}^i_{\hat{Q}}}(t)) + \eta_{\mathcal{I}^i_{\hat{Q}}}(t+1). \end{bmatrix} \tag{40}$$

Let $\mu = A_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t) + B_{\mathcal{I}^i_{\widehat{Q}}} u_{\mathcal{I}^i_{\widehat{Q}}}(t)$, $C = \begin{bmatrix} \mathbb{I} & 0 \\ K'_{\mathcal{I}^i_{\widehat{Q}}} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \Sigma_i^{\hat{x}\frac{1}{2}} & 0 \\ 0 & \Sigma_i^{\hat{u}\frac{1}{2}} \end{bmatrix}$, $g = \begin{bmatrix} w_{\mathcal{I}^i_{\widehat{Q}}}(t)) \\ \eta_{\mathcal{I}^i_{\widehat{Q}}}(t+1)) \end{bmatrix}$, and $Y = \text{smat}(y)$. Therefore, we can express

$$\begin{aligned} \langle \phi_{t+1}, y \rangle &= \begin{bmatrix} x_{t+1} \\ u_{t+1} \end{bmatrix}^{\intercal} Y \begin{bmatrix} x_{t+1} \\ u_{t+1} \end{bmatrix} \\ &= (\mu + Cg)^{\intercal} Y (\mu + Cg). \end{aligned} \tag{41}$$

is a Gaussian polynomial of degree 2.

Lemma F.1 (Gaussian hyper-contractivity (Proposition 5.48, Aubrun and Szarek (2017))) *For a Gaussian polynomial x of degree atmost k, $\forall\, q \geq 2$,*

$$\mathbb{E}[||x||_{L_q}] \leq (q-1)^{\frac{k}{2}} \mathbb{E}[||x||_{L_2}].$$

Therefore, from Lemma F.1 we have

$$\begin{aligned} &\mathbb{E}[|Z_{t+1}|_{L_4} | \mathcal{F}_t] \leq 3\mathbb{E}[|Z_{t+1}|_{L_2} | \mathcal{F}_t] \\ \Rightarrow\ &\mathbb{E}[|Z_{t+1}|^4 | \mathcal{F}_t] \leq 81 \mathbb{E}[|Z_{t+1}|^2 | \mathcal{F}_t]^2. \end{aligned} \tag{42}$$

Employing the Payley-Zygmund inequality, we conclude that for any $\theta \in (0,1)$,

$$\begin{aligned} P(|Z_{t+1}| \geq \sqrt{\theta \mathbb{E}[|Z_{t+1}|^2 | \mathcal{F}_t]} | \mathcal{F}_t) &= P(|Z_{t+1}|^2 \geq \theta \mathbb{E}[|Z_{t+1}|^2 | \mathcal{F}_t] | \mathcal{F}_t) \\ &\geq (1-\theta)^2 \frac{\mathbb{E}[|Z_{t+1}|^2 | \mathcal{F}_t]^2}{\mathbb{E}[|Z_{t+1}|^4 | \mathcal{F}_t]} \geq \frac{(1-\theta)^2}{81} \text{ (From (42))}. \end{aligned} \tag{43}$$

From (41), we have that

$$\begin{aligned} \mathbb{E}[Z] = \mathbb{E}[(\mu + Cg)^{\intercal} Y (\mu + Cg)] &= \mathbb{E}[\mu^{\intercal} Y \mu + g^{\intercal} C^{\intercal} Y C g + 2\mu^{\intercal} Y C g] \\ &= \mu^{\intercal} Y \mu + \text{Tr}(C^{\intercal} Y C) \ (\because g \text{ is zero mean}). \end{aligned} \tag{44}$$

Therefore,

$$Z - \mathbb{E}[Z] = g^{\intercal} C^{\intercal} Y C g + 2\mu^{\intercal} Y C g - \text{Tr}(C^{\intercal} Y C) \tag{45}$$

Assuming $\mathbb{E}(Z^2) \geq 2||C^{\intercal} Y C||_F^2$, we have that

$$\mathbb{E}[Z_{t+1}^2 | \mathcal{F}_t] \geq 2||C^{\intercal} Y C||_F^2 = 2||(C^{\intercal} \otimes C^{\intercal}) y|| \geq 2\sigma_{\min}^2 (C^{\intercal} \otimes C^{\intercal}) = 2\sigma_{\min}^4(C).$$

Substituting $\theta = \frac{1}{2}$ in (43) yields

$$P(|Z_{t+1}| \geq \sigma_{\min}^2(C) | \mathcal{F}_t) \geq \frac{1}{324}, \tag{46}$$

which establishes that $(Z_t)_{t \geq 1}$ satisfies the $(1, \sigma_{\min}^2(C), \frac{1}{324})$ block martingale small ball condition (BMSB) in Simchowitz et al. (2018).

To compute a crude upper bound on $||\mathbf{\Phi}||$ using Markov's inequality as

$$P(||\mathbf{\Phi}||^2 \geq t^2) \leq \frac{\mathbb{E}(||\mathbf{\Phi}||^2)}{t^2} = \frac{\mathbb{E}(\lambda_{\max}(\mathbf{\Phi}^\intercal\mathbf{\Phi}))}{t^2} \overset{(a)}{\leq} \frac{\mathbb{E}(\text{Tr}(\mathbf{\Phi}^\intercal\mathbf{\Phi}))}{t^2} \overset{(b)}{\leq} \frac{\text{Tr}(\mathbb{E}(\mathbf{\Phi}^\intercal\mathbf{\Phi}))}{t^2}, \tag{47}$$

where (a) follows from the fact that the $||\cdot||_2 \leq ||\cdot||_F$, and (b) is due to Jensen's inequality.

Now we upper bound $\mathbb{E}[||\phi_t||^2]$. Letting $z_t = (x_{\mathcal{I}^i_{\widehat{Q}}}(t), u_{\mathcal{I}^i_{\widehat{Q}}}(t))$, then we have $\mathbb{E}[||\phi_t||^2] = \mathbb{E}[||z_t||^4]$ By assumption, x_0 is zero mean which implies z_t is zero mean $\forall\, t$. Hence, $\mathbb{E}[||\phi_t||^2] = \mathbb{E}[||z_t||^4] \leq 3(\mathbb{E}[||z_t||^2])^2$.

Note that

$$\begin{aligned}
\mathbb{E}[||z_t||^2] &= \mathbb{E}[[x^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \;\; x^\intercal_{\mathcal{I}^i_{\widehat{Q}}}(t)K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}} + \eta^\intercal_{\mathcal{I}^i_{\widehat{Q}}}][x^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \;\; x^\intercal_{\mathcal{I}^i_{\widehat{Q}}}(t)K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}} + \eta^\intercal_{\mathcal{I}^i_{\widehat{Q}}}]^\intercal] \\
&= \mathbb{E}[x^\intercal_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}] + \mathbb{E}[x^\intercal_{\mathcal{I}^i_{\widehat{Q}}}(t)K'_{\mathcal{I}^i_{\widehat{Q}}} K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t)] + \mathbb{E}[x^\intercal_{\mathcal{I}^i_{\widehat{Q}}}(t)\eta_{\mathcal{I}^i_{\widehat{Q}}}] + \mathbb{E}[\eta^\intercal_{\mathcal{I}^i_{\widehat{Q}}} K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}} x_{\mathcal{I}^i_{\widehat{Q}}}(t)] + \mathbb{E}[\eta^\intercal_{\mathcal{I}^i_{\widehat{Q}}}\eta_{\mathcal{I}^i_{\widehat{Q}}}] \\
&= (1 + ||K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}}||^2)\text{Tr}(\mathbb{E}[x_{\mathcal{I}^i_{\widehat{Q}}} x^\intercal_{\mathcal{I}^i_{\widehat{Q}}}]) + \text{Tr}(\mathbb{E}[\eta_{\mathcal{I}^i_{\widehat{Q}}}\eta^\intercal_{\mathcal{I}^i_{\widehat{Q}}}]) \;(\because\; \mathbb{E}\eta = 0) \\
&= (1 + ||K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}}||^2)\text{Tr}(\Sigma^{\widehat{x}}_i) + \text{Tr}(\Sigma^{\widehat{u}}_i)
\end{aligned} \tag{48}$$

Assume that the initial global state is sampled from a zero-mean Gaussian distribution i.e., $x(0) \sim \mathcal{N}(0, \Sigma^x_0)$, which implies that $\forall\, t$, $\mathbb{E}(x(t)) = \mathbb{E}(u(t)) = 0$. Consider the propagation of the covariance

$$\begin{aligned}
\Sigma^{\widehat{x}}_i(t) &= \mathbb{E}(\widehat{x}^i(t)(\widehat{x}^i(t))^\intercal) \\
&= \mathbb{E}[(\widehat{A}^i_{t-1}\widehat{x}^i_{t-1} + \widehat{B}^i_{t-1}\widehat{u}^i_{t-1} + \widehat{w}^i_{t-1})(\widehat{A}^i_{t-1}\widehat{x}^i_{t-1} + \widehat{B}^i_{t-1}\widehat{u}^i_{t-1} + \widehat{w}^i_{t-1})^\intercal] \\
&= \mathbb{E}[((\widehat{A}^i_{t-1} + \widehat{B}^i_{t-1}\widehat{K}^i_{t-1})\widehat{x}^i_{t-1} + \widehat{B}^i_{t-1}\widehat{v}^i_{t-1} + \widehat{w}^i_{t-1}) \\
&\qquad\qquad ((\widehat{A}^i_{t-1} + \widehat{B}^i_{t-1}\widehat{K}^i_{t-1})\widehat{x}^i_{t-1} + \widehat{B}^i_{t-1}\widehat{v}^i_{t-1} + \widehat{w}^i_{t-1})^\intercal] \\
&= (\widehat{A}^i_{t-1} + \widehat{B}^i_{t-1}\widehat{K}^i_{t-1})\mathbb{E}[\widehat{x}^i_{t-1}\widehat{x}^{i\intercal}_{t-1}](\widehat{A}^i_{t-1} + \widehat{B}^i_{t-1}\widehat{K}^i_{t-1})^\intercal + \widehat{B}^i_{t-1}\mathbb{E}[\widehat{v}^i_{t-1}\widehat{v}^{i\intercal}_{t-1}]\widehat{B}^{i\intercal}_{t-1} \\
&\qquad\qquad + \mathbb{E}[\widehat{w}^i_{t-1}\widehat{w}^{i\intercal}_{t-1}].
\end{aligned} \tag{49}$$

Assuming time-invariant system and control gain matrices, recursive expansion of (49) yields

$$\begin{aligned}
\Sigma^{\widehat{x}}_i(t) &= (\widehat{A}^i + \widehat{B}^i\widehat{K}^i)^t\Sigma^{\widehat{x}}_i(0)((\widehat{A}^i + \widehat{B}^i\widehat{K}^i)^\intercal)^t \\
&\quad + \sum_{k=1}^{t}(\widehat{A}^i + \widehat{B}^i\widehat{K}^i)^{k-1}[\Sigma^{\widehat{x}}_w + \widehat{B}^i\Sigma^{\widehat{u}}_\eta\widehat{B}^{i\intercal}]((\widehat{A}^i + \widehat{B}^i\widehat{K}^i)^\intercal)^{k-1}.
\end{aligned} \tag{50}$$

Applying the trace operator on both sides yields

$$\begin{aligned}
\text{Tr}\left(\Sigma^{\widehat{x}}_i(t)\right) &= \text{Tr}\left((\widehat{A}^i + \widehat{B}^i\widehat{K}^i)^t\Sigma^{\widehat{x}}_i(0)((\widehat{A}^i + \widehat{B}^i\widehat{K}^i)^\intercal)^t\right) \\
&\quad + \text{Tr}\left(\sum_{k=1}^{t}(\widehat{A}^i + \widehat{B}^i\widehat{K}^i)^{k-1}[\Sigma^{\widehat{x}}_w + \widehat{B}^i\Sigma^{\widehat{u}}_\eta\widehat{B}^{i\intercal}]((\widehat{A}^i + \widehat{B}^i\widehat{K}^i)^\intercal)^{k-1}\right) \\
&\overset{(a)}{\leq} (n_{\widehat{x}} + n_{\widehat{u}})||\Sigma^{\widehat{x}}_i(0)||||(\widehat{A}^i + \widehat{B}^i\widehat{K}^i)^t||^2 \\
&\quad + \text{Tr}\left(\sum_{k=1}^{t}(\widehat{A}^i + \widehat{B}^i\widehat{K}^i)^{k-1}[\Sigma^{\widehat{x}}_w + \widehat{B}^i\Sigma^{\widehat{u}}_\eta\widehat{B}^{i\intercal}]((\widehat{A}^i + \widehat{B}^i\widehat{K}^i)^\intercal)^{k-1}\right),
\end{aligned} \tag{51}$$

where (a) follows from the identity that $\forall\ M \in \mathbb{R}^n, \|M\|_F \leq \sqrt{n}\,\|M\|$.

Assume that $(\widehat{A}^i + \widehat{B}^i \widehat{K}^i)$ is $(\tau, \rho)-$stable and $\widehat{K}^i$ stabilizes $(\widehat{A}^i, \widehat{B}^i)$. Let $\mathfrak{P}_\infty$ denote the unique solution of the Lyapunov equation

$$\mathfrak{P} = (\widehat{A}^i + \widehat{B}^i \widehat{K}^i)\mathfrak{P}(\widehat{A}^i + \widehat{B}^i \widehat{K}^i)^\intercal + \Sigma_w^{\widehat{x}} + \widehat{B}^i \Sigma_\eta^{\widehat{u}} \widehat{B}^{i\intercal}. \tag{52}$$

Then, (51) can be rewritten as

$$\mathrm{Tr}\left(\Sigma_i^{\widehat{x}}(t)\right) \leq (n_{\widehat{x}} + n_{\widehat{u}})||\Sigma_i^{\widehat{x}}(0)||(\tau\rho)^2 + \mathrm{Tr}\left(\mathfrak{P}_\infty\right), \tag{53}$$

Therefore, we have that

$$\begin{aligned}
\sqrt{\mathbb{E}[||\phi_t||^2]} &\leq \sqrt{3}((1 + ||K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}}||^2)\mathrm{Tr}(\Sigma_i^{\widehat{x}}(t)) + \mathrm{Tr}(\Sigma_i^{\widehat{u}}(t))) \\
&\text{Assuming } A_{\mathcal{I}^i_{\widehat{Q}}} + B_{\mathcal{I}^i_{\widehat{Q}}} K'_{\mathcal{I}^i_{\widehat{Q}}} \text{ is } (\tau, \rho)-\text{stable.} \\
&\leq \sqrt{3}((1 + ||K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\widehat{x}}||\Sigma_i^{\widehat{x}}(0)|| + \mathrm{Tr}(\mathfrak{P}_\infty)) + \mathrm{Tr}(\Sigma_i^{\widehat{u}})) \\
&\text{Assuming } \mathrm{Tr}(\Sigma_i^{\widehat{u}})) \leq \mathrm{Tr}(\Sigma_i^{\widehat{x}}(t)) \leq \mathrm{Tr}(\mathfrak{P}_\infty), \text{ we have} \\
&\leq 2\sqrt{3}((1 + ||K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\widehat{x}}||\Sigma_i^{\widehat{x}}(0)|| + \mathrm{Tr}(\mathfrak{P}_\infty)))
\end{aligned} \tag{54}$$

Therefore, the inequality in (47) can be written as

$$P(||\mathbf{\Phi}||^2 \geq t^2) \leq \frac{T(2\sqrt{3}((1 + ||K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\widehat{x}}||\Sigma_i^{\widehat{x}}(0)|| + \mathrm{Tr}(\mathfrak{P}_\infty))))^2}{t^2} \tag{55}$$

Choose $t^2 = \dfrac{T(2\sqrt{3}((1 + ||K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\widehat{x}}||\Sigma_i^{\widehat{x}}(0)|| + \mathrm{Tr}(\mathfrak{P}_\infty))))^2}{\delta}$, for some $\delta > 0$, then we have that

$$P\left(||\mathbf{\Phi}|| \geq \frac{2\sqrt{T}}{\sqrt{\delta}}\sqrt{3}((1 + ||K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\widehat{x}}||\Sigma_i^{\widehat{x}}(0)|| + \mathrm{Tr}(\mathfrak{P}_\infty)))\right) \leq \delta. \tag{56}$$

For some $\epsilon > 0$, let $\mathcal{N}(\epsilon)$ denote the $\epsilon-$net of the unit sphere $\mathcal{S}^{(n_{\widehat{x}}+n_{\widehat{u}})(n_{\widehat{x}}+n_{\widehat{u}}+1)/2-1}$. By Proposition 2.5 in Simchowitz et al. (2018), and (46), each $\nu \in \mathcal{N}(\epsilon)$ satisfies

$$P\left(||\mathbf{\Phi}\nu|| \leq \frac{\sigma^2_{\min}(C)\sqrt{T}}{324\sqrt{8}}\right) \leq e^{\frac{-T}{324^2 \cdot 8}} \tag{57}$$

From Corollary 4.2.13 in Vershynin (2018), we have that $\forall\ \epsilon > 0$, the covering number of the unit sphere S^{d-1} satisfies

$$N(\epsilon) \leq (1 + \frac{2}{\epsilon})^d.$$

Combining this with a union bound over all possible $\nu \in \mathcal{N}(\epsilon)$ yields

$$\begin{aligned}
P\left(\min_{\nu \in \mathcal{N}(\epsilon)} ||\mathbf{\Phi}\nu|| \geq \frac{\sigma^2_{\min}(C)\sqrt{T}}{324\sqrt{8}}\right) &\geq 1 - (1 + 2/\epsilon)^{(n_{\widehat{x}}+n_{\widehat{u}})(n_{\widehat{x}}+n_{\widehat{u}}+1)/2} e^{\frac{-T}{324^2 \cdot 8}} \\
&\geq 1 - (1 + 2/\epsilon)^{(n_{\widehat{x}}+n_{\widehat{u}})^2} e^{\frac{-T}{324^2 \cdot 8}}.
\end{aligned} \tag{58}$$

We wish to choose $\delta,\ \epsilon$ such that $P\left(\min_{\nu\in\mathcal{N}(\epsilon)}||\mathbf{\Phi}\nu|| \geq \frac{\sigma^2_{\text{min}}(C)\sqrt{T}}{324\sqrt{8}}\right) \geq 1-\frac{\delta}{2}$, which implies

$$\begin{aligned}
&(1+2/\epsilon)^{(n_{\widehat{x}}+n_{\widehat{u}})^2}e^{\frac{-T}{324^2\cdot 8}} = \frac{\delta}{2} \\
&\Rightarrow \epsilon = \frac{2}{\exp\left(\frac{1}{(n_{\widehat{x}}+n_{\widehat{u}})^2}\log\left(\frac{\delta}{2}\right)+\frac{T}{(n_{\widehat{x}}+n_{\widehat{u}})^2\cdot 324^2\cdot 8}\right)-1}.
\end{aligned} \tag{59}$$

Note that

$$\sigma_{\text{min}}(\mathbf{\Phi}) = \inf_{||\nu||=1}||\mathbf{\Phi}\nu|| \geq \min_{\nu\in\mathcal{N}(\epsilon)}||\mathbf{\Phi}\nu|| - ||\mathbf{\Phi}||\epsilon. \tag{60}$$

Since,

$$P\left(||\mathbf{\Phi}|| \geq \frac{2\sqrt{T}}{\sqrt{\delta}}\sqrt{6}((1+||K'^{\intercal}_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\widehat{x}}||\Sigma^{\widehat{x}}_i(0)||+\text{Tr}(\mathfrak{P}_\infty)))\right) \leq \frac{\delta}{2},$$

$$P\left(\min_{\nu\in\mathcal{N}(\epsilon)}||\mathbf{\Phi}\nu|| \leq \frac{\sigma^2_{\text{min}}(C)\sqrt{T}}{324\sqrt{8}}\right) \leq \frac{\delta}{2},$$

taking a union bound yields that with probability $1-\delta$,

$$\sigma_{\text{min}}(\mathbf{\Phi}) \geq \frac{\sigma^2_{\text{min}}(C)\sqrt{T}}{324\sqrt{8}} - \frac{2\epsilon\sqrt{T}}{\sqrt{\delta}}\sqrt{6}((1+||K'^{\intercal}_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\widehat{x}}||\Sigma^{\widehat{x}}_i(0)||+\text{Tr}(\mathfrak{P}_\infty))). \tag{61}$$

Substituting (59) in (61) and setting the RHS to be non-negative yields

$$\begin{aligned}
&\frac{\sigma^2_{\text{min}}(C)\sqrt{T}}{324\sqrt{8}} - \frac{2\epsilon\sqrt{T}}{\sqrt{\delta}}\sqrt{6}((1+||K'^{\intercal}_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\widehat{x}}||\Sigma^{\widehat{x}}_i(0)||+\text{Tr}(\mathfrak{P}_\infty))) \geq 0 \\
&\frac{\sigma^2_{\text{min}}(C)}{324\sqrt{8}} \geq \frac{2\cdot 2\cdot\sqrt{6}}{\sqrt{\delta}}\frac{((1+||K'^{\intercal}_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\widehat{x}}||\Sigma^{\widehat{x}}_i(0)||+\text{Tr}(\mathfrak{P}_\infty)))}{\exp\left(\frac{1}{(n_{\widehat{x}}+n_{\widehat{u}})^2}\log\left(\frac{\delta}{2}\right)+\frac{T}{(n_{\widehat{x}}+n_{\widehat{u}})^2\cdot 324^2\cdot 8}\right)-1} \\
&\frac{1}{(n_{\widehat{x}}+n_{\widehat{u}})^2}\log\left(\frac{\delta}{2}\right)+\frac{T}{(n_{\widehat{x}}+n_{\widehat{u}})^2\cdot 324^2\cdot 8} \geq \\
&\log\left(1+\frac{2\cdot 2\cdot\sqrt{6}\cdot 324\cdot\sqrt{8}}{\sqrt{\delta}}\frac{((1+||K'^{\intercal}_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\widehat{x}}||\Sigma^{\widehat{x}}_i(0)||+\text{Tr}(\mathfrak{P}_\infty)))}{\sigma^2_{\text{min}}(C)}\right) \\
&T \geq 324^2\cdot 8\left((n_{\widehat{x}}+n_{\widehat{u}})^2\log\left(1+\frac{10368\sqrt{3}}{\sqrt{\delta}}\frac{(1+||K'^{\intercal}_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\widehat{x}}||\Sigma^{\widehat{x}}_i(0)||+\text{Tr}(\mathfrak{P}_\infty))}{\sigma^2_{\text{min}}(C)}\right)+\log(\frac{2}{\delta})\right)
\end{aligned} \tag{62}$$

Substituting (62) in (59) yields

$$\epsilon = \frac{\sqrt{\delta}}{5184\sqrt{3}}\frac{\sigma^2_{\text{min}}(C)}{(1+||K'^{\intercal}_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\widehat{x}}||\Sigma^{\widehat{x}}_i(0)||+\text{Tr}(\mathfrak{P}_\infty))}. \tag{63}$$

For the choice of ϵ according to (63) we can ensure that with probability $1-\delta$,

$$\sigma_{\min}(\mathbf{\Phi}) \geq \frac{\sigma^2_{\min}(C)\sqrt{T}}{324\sqrt{8}} - \frac{2\sigma^2_{\min}(C)\cdot\sqrt{2}\sqrt{T}}{2592} = \frac{\sigma^2_{\min}(C)\sqrt{T}}{648\sqrt{8}} \tag{64}$$

Note that if $\Sigma_i^{\hat{u}} = \sigma_\eta^2\mathbb{I},\ \Sigma_i^{\hat{x}} = \sigma_w^2\mathbb{I}$, and $\sigma_\eta \leq \sigma_w$ then

$$\begin{aligned}
\sigma^2_{\min}(C) = \lambda_{\min}(C^\intercal C) = \lambda_{\min}(CC^\intercal) &= \lambda_{\min}\left(\begin{bmatrix} {\Sigma_i^{\hat{x}}}^{\frac{1}{2}} & 0 \\ K'_{\mathcal{I}^i_{\widehat{Q}}}{\Sigma_i^{\hat{x}}}^{\frac{1}{2}} & {\Sigma_i^{\hat{u}}}^{\frac{1}{2}} \end{bmatrix}\begin{bmatrix} {\Sigma_i^{\hat{x}}}^{\frac{1}{2}} & {\Sigma_i^{\hat{x}}}^{\frac{1}{2}}K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}} \\ 0 & {\Sigma_i^{\hat{u}}}^{\frac{1}{2}} \end{bmatrix}\right) \\
&\overset{(a)}{\geq} \frac{\sigma_\eta^2\lambda_{\min}(\sigma_w^2\mathbb{I})}{2\sigma_w^2||K'_{\mathcal{I}^i_{\widehat{Q}}}K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}}||_2 + \sigma_\eta^2} \\
&\overset{(b)}{\geq} \frac{\sigma_\eta^2\sigma_w^2}{2\sigma_w^2||K'_{\mathcal{I}^i_{\widehat{Q}}}K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}}||_2 + 2\sigma_w^2} \\
&= \frac{\sigma_\eta^2}{2(1+||K'_{\mathcal{I}^i_{\widehat{Q}}}||^2)},
\end{aligned} \tag{65}$$

where (a) follows from Lemma F.6 in Dean et al. (2018) and (b) is due to the assumption that $\sigma_\eta \leq \sigma_w$.

Therefore, by combining (64) and (65), we conclude that if $A_{\mathcal{I}^i_{\widehat{Q}}} + B_{\mathcal{I}^i_{\widehat{Q}}}K'_{\mathcal{I}^i_{\widehat{Q}}}$ is $(\tau,\rho)-$stable, $\Sigma_i^{\hat{u}} = \sigma_\eta^2\mathbb{I},\ \Sigma_i^{\hat{x}} = \sigma_w^2\mathbb{I}$, and $\sigma_\eta \leq \sigma_w$, and

$$T \geq 324^2{\cdot}8\left((n_{\hat{x}}+n_{\hat{u}})^2\log\left(1+\frac{10368\sqrt{3}}{\sqrt{\delta}}\frac{(1+||K'^\intercal_{\mathcal{I}^i_{\widehat{Q}}}||^2)(\tau^2\rho^2 n_{\hat{x}}||\Sigma_i^{\widehat{x}}(0)||+\text{Tr}(\mathfrak{P}_\infty))}{\sigma^2_{\min}(C)}\right)+\log(\frac{2}{\delta})\right),$$

then for some $\delta > 0$, with probability $1-\delta$,

$$\sigma_{\min}(\mathbf{\Phi}) \geq \frac{\sigma_\eta^2\sqrt{T}}{1296\sqrt{8}(1+||K'_{\mathcal{I}^i_{\widehat{Q}}}||^2)}.$$

F.2. Upper bound of the numerator in (39)

For the remainder of the proof, let $c_i\ \forall\ i$ denote universal constants. Define

$$V_1 := c_1\frac{\sigma_\eta^4}{(1+||K'_{\mathcal{I}^i_{\widehat{Q}}}||^2)^2}T\cdot\mathbb{I},$$

$$V_2 := c_2\frac{T}{\delta}(1+||K'_{\mathcal{I}^i_{\widehat{Q}}}||^2)^2(\tau^2\rho^2 n||\Sigma_0||+\text{Tr}(\mathfrak{P}_\infty))^2\cdot\mathbb{I}.$$

Then, by the above proposition and (56), define an event ζ_1 such that with atleast probability $1-\delta$ we have that

$$V_1 \preceq \mathbf{\Phi}^\intercal\mathbf{\Phi} \preceq V_2.$$

For brevity, let $x_{\mathcal{I}^i_{\widehat{Q}}}(t) = \widehat{x}^i_t,\ u_{\mathcal{I}^i_{\widehat{Q}}}(t) = \widehat{u}^i_t,\ A_{\mathcal{I}^i_{\widehat{Q}}}(t) = \widehat{A}^i_t,\ B_{\mathcal{I}^i_{\widehat{Q}}}(t) = \widehat{B}^i_t,$ and $K_{\mathcal{I}^i_{\widehat{Q}}}(t) = \widehat{K}^i_t$. Note that,

$$\begin{aligned}\mathbb{E}[\widehat{x}^i_{t+1}\widehat{x}^{i\intercal}_{t+1}|\widehat{x}^i_t,\widehat{u}^i_t] - \widehat{x}^i_{t+1}\widehat{x}^{i\intercal}_{t+1} &= \mathbb{E}[(\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t + \widehat{w}^i_t)(\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t + \widehat{w}^i_t)^\intercal|\widehat{x}^i_t,\widehat{u}^i_t] \\ &\quad - (\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t + \widehat{w}^i_t)(\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t + \widehat{w}^i_t)^\intercal, \\ &= (\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t)(\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t)^\intercal + \widehat{\Sigma^x_{wt}}^i - (\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t)(\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t)^\intercal \\ &\quad - \widehat{w}^i_t(\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t)^\intercal - (\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t)\widehat{w}^{i\intercal}_t - \widehat{w}^i_t\widehat{w}^{i\intercal}_t. \end{aligned} \tag{66}$$

Therefore, it is straightforward to verify that

$$\begin{aligned}&\mathbb{E}[\psi_{t+1}|\widehat{x}^i_t,\widehat{u}^i_t] - \psi_{t+1} = \\ &\mathrm{svec}\left(\begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}(\sigma^2_w\mathbb{I} - \widehat{w}^i_t(\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t)^\intercal - (\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t)\widehat{w}^{i\intercal}_t - \widehat{w}^i_t\widehat{w}^{i\intercal}_t)\begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}^\intercal\right).\end{aligned} \tag{67}$$

Taking inner product of (67) with p yields

$$\begin{aligned}&(\mathbb{E}[\psi_{t+1}|\widehat{x}^i_t,\widehat{u}^i_t] - \psi_{t+1})^\intercal p = \\ &\mathrm{Tr}\left((\sigma^2_w\mathbb{I} - \widehat{w}^i_t(\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t)^\intercal - (\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t)\widehat{w}^{i\intercal}_t - \widehat{w}^i_t\widehat{w}^{i\intercal}_t)\begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}\right) \\ &= \mathrm{Tr}\left((\sigma^2_w\mathbb{I} - \widehat{w}^i_t\widehat{w}^{i\intercal}_t)\begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}\right) - 2\widehat{w}^{i\intercal}_t\begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}(\widehat{A}^i_t\widehat{x}^i_t + \widehat{B}^i_t\widehat{u}^i_t)\end{aligned} \tag{68}$$

Theorem F.1 (Hanson-Wright Inequality (Theorem 6.2.1, Vershynin (2018))) *Let $X = (X1, \cdots, Xn) \in \mathbb{R}^n$ be a random vector with independent, mean zero, sub-Gaussian coordinates. Let A be a $n \times n$ matrix, and $K = max_i||X_i||_{\psi_2}$. Then, $\forall\ t > 0$, we have*

$$P(|X^\intercal AX - \mathbb{E}X^\intercal AX| \geq t) \leq 2\exp\left(-c\ min\left(\frac{t^2}{K^4||A||^2_F}, \frac{t}{K^2||A||_{op}}\right)\right),$$

Let

$$\begin{aligned}\delta/T = 2\exp\left(-c\frac{t^2}{K^4||A||^2_F}\right) &\Rightarrow t^2 = \frac{K^4||A||^2_F}{c}\log(\frac{2T}{\delta}),\ \text{(or)} \\ \delta/T = 2\exp\left(-c\frac{t}{K^2||A||_{\text{op}}}\right) &\Rightarrow t = \frac{K^2||A||_{\text{op}}}{c}\log(\frac{2T}{\delta})\end{aligned} \tag{69}$$

From (69), (68) we have that with probability $1 - \frac{\delta}{T}$,

$$\begin{aligned}&\left|\mathrm{Tr}\left((\sigma^2_w\mathbb{I} - \widehat{w}^i_t\widehat{w}^{i\intercal}_t)\begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}\right)\right| \leq \\ &\min\left(\frac{K^2\left\|\begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}\right\|}{c}\log(\frac{2T}{\delta}), \frac{K^2\left\|\begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}\right\|_F}{\sqrt{c}}\sqrt{\log(\frac{2T}{\delta})}\right).\end{aligned} \tag{70}$$

Since, $\widehat{w}_t^i \sim \mathcal{N}(0, \sigma_w^2 \mathbb{I})$, $K = \sigma_w$, and $\left\| \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \right\| \leq \left\| \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal \right\| ||P_i|| \left\| \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \right\|$.

Note that

$$\left\| \begin{bmatrix} \mathbb{I} \\ \widehat{K} \end{bmatrix} \right\| = \sqrt{\lambda_{\max}(\mathbb{I} + \widehat{K}^\intercal \widehat{K})} = \sqrt{\max_{x \in S_{n_{\widehat{x}}}} \frac{x^\intercal (\mathbb{I} + \widehat{K}^\intercal \widehat{K}) x}{x^\intercal x}} = \sqrt{1 + \max_{x \in S_{n_{\widehat{x}}}} \frac{x^\intercal \widehat{K}^\intercal \widehat{K} x}{x^\intercal x}} = \sqrt{1 + \lambda_{\max}(\widehat{K}^\intercal \widehat{K})}. \tag{71}$$

Since, $||M|| = ||M^\intercal||$ and $||M|| \leq ||M||_F$, we have that

$$||A|| \leq (1 + \lambda_{\max}(\widehat{K}^\intercal \widehat{K}))||P_i|| \leq (1 + ||\widehat{K}||^2)||P_i||_F, \tag{72}$$

Similarly, we have

$$||A||_F \leq (\sqrt{n_{\widehat{x}}})||A|| \leq (\sqrt{n_{\widehat{x}}})(1 + ||\widehat{K}||^2)||P_i||_F. \tag{73}$$

Therefore, we have that

$$\begin{aligned}
&\left| \text{Tr} \left((\sigma_w^2 \mathbb{I} - \widehat{w}_t^i \widehat{w}_t^{i\intercal}) \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \right) \right| \leq \\
&\min \left(\frac{\sigma_w^2 (1 + ||\widehat{K}||^2)||P_i||_F}{c} \log(\frac{2T}{\delta}), \frac{\sigma_w^2 (\sqrt{n_{\widehat{x}}})(1 + ||\widehat{K}||^2)||P_i||_F.}{\sqrt{c}} \sqrt{\log(\frac{2T}{\delta})} \right) \\
\Rightarrow &\left| \text{Tr} \left((\sigma_w^2 \mathbb{I} - \widehat{w}_t^i \widehat{w}_t^{i\intercal}) \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \right) \right| \leq c_1 \sigma_w^2 (1 + ||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F \log(\frac{T}{\delta}),
\end{aligned} \tag{74}$$

where c_1 is a universal constant.

From proposition 4.7 in Tu and Recht (2018), with probability $1 - \delta/T$,

$$\begin{aligned}
&\left| w_t^\intercal \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} (Ax_t + Bu_t) \right| \leq \\
&\min \left(\frac{\left\| \Sigma_w^{\widehat{x}\frac{1}{2}} \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \Sigma_{t+1}^{\widehat{x}\frac{1}{2}} \right\|_F}{\sqrt{c}} \sqrt{\log(2T/\delta)}, \frac{\left\| \Sigma_w^{\widehat{x}\frac{1}{2}} \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \Sigma_{t+1}^{\widehat{x}\frac{1}{2}} \right\|}{c} \log(2T/\delta) \right)
\end{aligned} \tag{75}$$

Let $L' = \widehat{A}^i + \widehat{B}^i \widehat{K}'^i$. Then, observe that

$$\begin{aligned}
&\left\| \Sigma_w^{\widehat{x}\frac{1}{2}} \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \Sigma_{t+1}^{\widehat{x}\frac{1}{2}} \right\|_F \leq \left\| \Sigma_w^{\widehat{x}\frac{1}{2}} \right\|_F \left\| \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \right\|_F \left\| \Sigma_{t+1}^{\widehat{x}\frac{1}{2}} \right\|_F \\
&\leq \|\sigma_w I\|_F \sqrt{n_{\widehat{x}}} (1 + ||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F \sqrt{\text{Tr}\left(L^{t+1} \Sigma_i^{\widehat{x}}(0) (L^\intercal)^{t+1}\right) + \text{Tr}\left(\mathfrak{P}_{t+1}\right)} \\
&\qquad \text{(From (73), (51))} \\
&= n_{\widehat{x}} \sigma_w (1 + ||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F \sqrt{\text{Tr}\left(L^{t+1} \Sigma_i^{\widehat{x}}(0) (L^\intercal)^{t+1}\right) + \text{Tr}\left(\mathfrak{P}_t\right) + \text{Tr}\left(\sigma_w^2 \mathbb{I} + \sigma_\eta^2 \widehat{B} \widehat{B}^\intercal\right)}.
\end{aligned} \tag{76}$$

Similarly, it is straightforward to show that

$$\left\|\Sigma_w^{\widehat{x}\frac{1}{2}}\begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}\Sigma_{t+1}^{\widehat{x}\frac{1}{2}}\right\| \leq \left\|\Sigma_w^{\widehat{x}\frac{1}{2}}\right\|\left\|\begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}\right\|\left\|\Sigma_{t+1}^{\widehat{x}\frac{1}{2}}\right\|$$

$$\leq \|\sigma_w I\|\,(1+||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F\sqrt{\left\|L^{t+1}\Sigma_i^{\widehat{x}}(0)(L^\intercal)^{t+1}+\mathfrak{P}_{t+1}\right\|}$$

(From (72), (51), and using $||M^{\frac{1}{2}}|| = ||M||^{\frac{1}{2}}$, $\forall$ symmetric, p.s.d M.)

$$= \sigma_w(1+||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F\sqrt{\left\|L^{t+1}\Sigma_i^{\widehat{x}}(0)(L^\intercal)^{t+1}+L\mathfrak{P}_t L^\intercal+\sigma_w^2\mathbb{I}+\sigma_\eta^2\widehat{B}\widehat{B}^\intercal\right\|}. \tag{77}$$

Therefore, combining (76), (77) yields

$$\left|w_t^\intercal\begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}(Ax_t+Bu_t)\right| \leq$$
$$\sigma_w(1+||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F$$
$$\min\left(\frac{n_{\widehat{x}}\sqrt{\mathrm{Tr}\left(L^{t+1}\Sigma_i^{\widehat{x}}(0)(L^\intercal)^{t+1}\right)+\mathrm{Tr}\left(L\mathfrak{P}_tL^\intercal\right)+\mathrm{Tr}\left(\sigma_w^2\mathbb{I}+\sigma_\eta^2\widehat{B}\widehat{B}^\intercal\right)}}{\sqrt{c}}\sqrt{\log(2T/\delta)},\right.$$
$$\left.\frac{\sqrt{\left\|L^{t+1}\Sigma_i^{\widehat{x}}(0)(L^\intercal)^{t+1}+L\mathfrak{P}_tL^\intercal+\sigma_w^2\mathbb{I}+\sigma_\eta^2\widehat{B}\widehat{B}^\intercal\right\|}}{c}\log(2T/\delta)\right) \tag{78}$$

Since, $||\cdot|| \leq ||\cdot||_F$, and min $(\log(2T/\delta)/c, \sqrt{\log(2T/\delta)/c}) \leq \log(2T/\delta)/c$. Thus, we have that

$$\left|w_t^\intercal\begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}^\intercal P_i \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}}\end{bmatrix}(Ax_t+Bu_t)\right| \leq \frac{\sigma_w(1+||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F}{c}\sqrt{\left\|L^{t+1}\Sigma_i^{\widehat{x}}(0)(L^\intercal)^{t+1}+L\mathfrak{P}_tL^\intercal+\sigma_w^2\mathbb{I}+\sigma_\eta^2\widehat{B}\widehat{B}^\intercal\right\|}$$
$$\leq c_1\sigma_w(1+||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F\sqrt{\left\|L^{t+1}\Sigma_i^{\widehat{x}}(0)(L^\intercal)^{t+1}\right\|+\|L\mathfrak{P}_tL^\intercal\|+\left\|\sigma_w^2\mathbb{I}+\sigma_\eta^2\widehat{B}\widehat{B}^\intercal\right\|}\log(2T/\delta)$$
$$\overset{(a)}{\leq} c_1\sigma_w(1+||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F\sqrt{\tau^2\rho^{2(t+1)}\left\|\Sigma_i^{\widehat{x}}(0)\right\|+\|\mathfrak{P}_\infty\|+\sigma_w^2+\left\|\sigma_\eta^2\widehat{B}\widehat{B}^\intercal\right\|}\log(2T/\delta), \tag{79}$$

where (a) is due to $\mathfrak{P}_t \preceq \mathfrak{P}_\infty$, $(\hat{A}^i+\hat{B}^i\hat{K}^i)\mathfrak{P}_\infty(\hat{A}^i+\hat{B}^i\hat{K}^i)^\intercal \preceq \mathfrak{P}_\infty$, and $(\hat{A}^i+\hat{B}^i\hat{K}^i)$ is $(\tau,\rho)-$stable.

From (68), (74), and (79) we obtain that

$$\begin{aligned}
&(\mathbb{E}[\psi_{t+1}|\widehat{x}_t^i, \widehat{u}_t^i] - \psi_{t+1})^\intercal p \leq \\
&c_2(\sigma_w^2 + \sigma_w\sqrt{\tau^2\rho^{2(t+1)}||\Sigma^{\widehat{x}}(0)|| + ||\mathfrak{P}_\infty|| + \sigma_w^2 + \sigma_\eta^2||B||^2})(1 + ||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F \log(T/\delta) \\
&\overset{(a)}{\leq} c_3'\sigma_w\sqrt{\tau^2\rho^{2(t+1)}||\Sigma^{\widehat{x}}(0)|| + ||\mathfrak{P}_\infty|| + \sigma_w^2 + \sigma_\eta^2||B||^2}(1 + ||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F \log(T/\delta) \\
&\leq c_3'\sigma_w\sqrt{\tau^2\rho^4||\Sigma^{\widehat{x}}(0)|| + ||\mathfrak{P}_\infty|| + \sigma_w^2 + \sigma_\eta^2||B||^2}(1 + ||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F \log(T/\delta), \ \forall\, t \geq 1 \\
&\leq \frac{\left(c_3\sigma_w\sqrt{\tau^2\rho^4||\Sigma^{\widehat{x}}(0)|| + ||\mathfrak{P}_\infty|| + \sigma_w^2 + \sigma_\eta^2||B||^2}(1 + ||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F \log(T/\delta)\right)^2}{2} \qquad (80)
\end{aligned}$$

where the inequality (a) holds because $\mathfrak{P}_\infty \succeq \sigma_w^2\mathbb{I}$ and hence $\sigma_w \leq ||\mathfrak{P}_\infty||^{\frac{1}{2}}$.
Therefore, $\forall\, \lambda \in \mathbb{R}$ and $t \geq 1$, (80) can be rewritten as

$$\begin{aligned}
&e^{\lambda(\mathbb{E}[\psi_{t+1}|\widehat{x}_t^i,\widehat{u}_t^i] - \psi_{t+1})^\intercal p)} \leq e^{\frac{\lambda^2R^2}{2}} \\
&\Rightarrow \mathbb{E}[e^{\lambda(\mathbb{E}[\psi_{t+1}|\widehat{x}_t^i,\widehat{u}_t^i] - \psi_{t+1})^\intercal p}] \leq e^{\frac{\lambda^2R^2}{2}}, \qquad (81)
\end{aligned}$$

where $R = c_3\sigma_w\sqrt{\tau^2\rho^4||\Sigma^{\widehat{x}}(0)|| + ||\mathfrak{P}_\infty|| + \sigma_w^2 + \sigma_\eta^2||B||^2}(1 + ||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)||P_i||_F \log(T/\delta)$. Therefore, for any fixed $v \in \mathbb{R}^{n_{\widehat{x}}+n_{\widehat{u}}}$, $(\mathbb{E}[\psi_{t+1}|\widehat{x}_t^i, \widehat{u}_t^i] - \psi_{t+1})^\intercal v$ is R-sub-Gaussian.

To upper bound $(\mathbf{\Phi}^\intercal\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\intercal(\mathbf{\Xi} - \mathbf{\Psi}_+)p$, note that as

$$V_1 \preceq (\mathbf{\Phi}^\intercal\mathbf{\Phi}) \Rightarrow V_1 + \mathbf{\Phi}^\intercal\mathbf{\Phi} \preceq 2(\mathbf{\Phi}^\intercal\mathbf{\Phi}) \Rightarrow (V_1 + \mathbf{\Phi}^\intercal\mathbf{\Phi})^{-1} \succeq \frac{1}{2}(\mathbf{\Phi}^\intercal\mathbf{\Phi})^{-1} \Rightarrow ||2(V_1 + \mathbf{\Phi}^\intercal\mathbf{\Phi})^{-1}|| \geq ||(\mathbf{\Phi}^\intercal\mathbf{\Phi})^{-1}||.$$

From Corollary 1 in Abbasi-Yadkori et al. (2011) we have that w.p $1 - \delta/2$,

$$\begin{aligned}
||(V_1 + (\mathbf{\Phi}^\intercal\mathbf{\Phi}))^{-\frac{1}{2}}\mathbf{\Phi}^\intercal(\mathbf{\Xi} - \mathbf{\Psi}_+)p|| &\leq ||(V_1 + (\mathbf{\Phi}^\intercal\mathbf{\Phi}))^{-\frac{1}{2}}\mathbf{\Phi}^\intercal(\mathbf{\Xi} - \mathbf{\Psi}_+)p|| \\
&\leq 2\sqrt{2}R\sqrt{\log\left(2 \cdot 5^{(n_{\widehat{x}}+n_{\widehat{u}})((n_{\widehat{x}}+n_{\widehat{u}}+1)/2} \cdot \frac{\det\left((V_1 + (\mathbf{\Phi}^\intercal\mathbf{\Phi}))^{\frac{1}{2}}V_1^{\frac{-1}{2}}\right)}{\delta}\right)} \\
&\leq \sqrt{2}R\left[\sqrt{\log\left(2 \cdot 5^{(n_{\widehat{x}}+n_{\widehat{u}})((n_{\widehat{x}}+n_{\widehat{u}}+1)/2}\right)} + \sqrt{\log\left(\frac{\det\left((V_1 + (\mathbf{\Phi}^\intercal\mathbf{\Phi}))V_1^{-1}\right)^{\frac{1}{2}}}{\delta}\right)}\right] \\
&\leq c_4R(n_{\widehat{x}} + n_{\widehat{u}}) + Rc_5\sqrt{\log\left(\frac{\det\left((V_1 + (\mathbf{\Phi}^\intercal\mathbf{\Phi}))V_1^{-1}\right)^{\frac{1}{2}}}{\delta}\right)}, \qquad (82)
\end{aligned}$$

where the last inequality follows from $(n_{\widehat{x}} + n_{\widehat{u}})((n_{\widehat{x}} + n_{\widehat{u}} + 1)/2 \leq (n_{\widehat{x}} + n_{\widehat{u}})^2$, Call this event $\mathcal{E}_2$.

For the remainder of the proof we consider the case where $\mathcal{E}_1 \cap \mathcal{E}_2$ occurs with probability $1-\delta$. Therefore, by application of Lemma A.5 in Krauth et al. (2019), we have that

$$
\begin{aligned}
&||(\mathbf{\Phi}^\intercal\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\intercal(\mathbf{\Xi}-\mathbf{\Psi}_+)p|| \leq \sqrt{2}\left\|(V_1+(\mathbf{\Phi}^\intercal\mathbf{\Phi}))^{-\frac{1}{2}}\mathbf{\Phi}^\intercal(\mathbf{\Xi}-\mathbf{\Psi}_+)p\right\| \\
&\leq c_6(n_{\widehat{x}}+n_{\widehat{u}})R + c_7 R\sqrt{\log\left(\frac{\det\left((V_1+(\mathbf{\Phi}^\intercal\mathbf{\Phi}))^{\frac{1}{2}}V_1^{-\frac{1}{2}}\right)}{\delta}\right)} \\
&\leq c_6(n_{\widehat{x}}+n_{\widehat{u}})R + c_7 R\sqrt{\log\left(\frac{\det\left((V_1+V_2)^{\frac{1}{2}}V_1^{\frac{-1}{2}}\right)}{\delta}\right)} \\
&= c_6(n_{\widehat{x}}+n_{\widehat{u}})R + c_7 R\sqrt{\log\left(\frac{\det\left((V_1+V_2)V_1^{-1}\right)^{\frac{1}{2}}}{\delta}\right)} \\
&= c_6(n_{\widehat{x}}+n_{\widehat{u}})R + c_7 R\sqrt{\log\left(\frac{\det\left((\mathbb{I}+c_1^{-1}c_2\frac{1}{\delta\sigma_\eta^4}(1+||K'_{\mathcal{I}^i_{\widehat{Q}}}||^2)^4(\tau^2\rho^2 n_{\widehat{x}}||\Sigma_0||+\mathrm{Tr}(\mathfrak{P}_\infty))^2\cdot\mathbb{I})\right)^{\frac{1}{2}}}{\delta}\right)} \\
&= c_6(n_{\widehat{x}}+n_{\widehat{u}})R + c_7 R\sqrt{\log\left(\left(\frac{1}{\delta^{\frac{2}{(n_{\widehat{x}}+n_{\widehat{u}})^2}}}(1+c_1^{-1}c_2\frac{1}{\delta\sigma_\eta^4}(1+||K'_{\mathcal{I}^i_{\widehat{Q}}}||^2)^4(\tau^2\rho^2 n_{\widehat{x}}||\Sigma_0||+\mathrm{Tr}(\mathfrak{P}_\infty))^2)\right)^{\frac{(n_{\widehat{x}}+n_{\widehat{u}})^2}{2}}\right)} \\
&= c_6(n_{\widehat{x}}+n_{\widehat{u}})R \\
&\quad + c_7 R\frac{(n_{\widehat{x}}+n_{\widehat{u}})}{\sqrt{2}}\sqrt{\frac{2}{(n_{\widehat{x}}+n_{\widehat{u}})^2}\log\left(\frac{1}{\delta}\right)+\log\left((1+c_1^{-1}c_2\frac{1}{\delta\sigma_\eta^4}(1+||K'_{\mathcal{I}^i_{\widehat{Q}}}||^2)^4(\tau^2\rho^2 n_{\widehat{x}}||\Sigma_0||+\mathrm{Tr}(\mathfrak{P}_\infty))^2)\right)} \\
&\leq c_6(n_{\widehat{x}}+n_{\widehat{u}})R \\
&\quad + c_7 R\frac{(n_{\widehat{x}}+n_{\widehat{u}})}{\sqrt{2}}\sqrt{\log\left(\frac{1}{\delta}\right)+\log\left((1+c_1^{-1}c_2\frac{1}{\delta\sigma_\eta^4}(1+||K'_{\mathcal{I}^i_{\widehat{Q}}}||^2)^4(\tau^2\rho^2 n_{\widehat{x}}||\Sigma_0||+\mathrm{Tr}(\mathfrak{P}_\infty))^2)\right)} \quad (\because n_{\widehat{x}}, n_{\widehat{u}} \geq 1) \\
&= c_6(n_{\widehat{x}}+n_{\widehat{u}})R + c_7\frac{(n_{\widehat{x}}+n_{\widehat{u}})}{\sqrt{2}}R\sqrt{\log\left(\frac{1}{\delta}+\frac{1}{\delta^2\sigma_\eta^4}(1+||K'_{\mathcal{I}^i_{\widehat{Q}}}||^2)^4(\tau^2\rho^2 n_{\widehat{x}}||\Sigma_0||+\mathrm{Tr}(\mathfrak{P}_\infty))^2\right)}
\end{aligned}
$$

Note that $\forall\, \zeta \in \mathbb{R}_{>0}$, $\exists$ a sufficiently large $M \in \mathbb{R}$, such that $\log\left(\frac{1}{\delta}+\frac{\zeta}{\delta^2}\right) \leq M^2\log\left(\frac{\zeta}{\delta}\right)$, Therefore, we have that

$$
\begin{aligned}
&\leq c_6(n_{\widehat{x}}+n_{\widehat{u}})R + c_8(n_{\widehat{x}}+n_{\widehat{u}})R\sqrt{\log\left(\frac{1}{\delta\sigma_\eta^4}(1+||K'_{\mathcal{I}^i_{\widehat{Q}}}||^2)^4(\tau^2\rho^2 n_{\widehat{x}}||\Sigma_0||+\mathrm{Tr}(\mathfrak{P}_\infty))^2\right)} \\
&\leq c(n_{\widehat{x}}+n_{\widehat{u}})R \;\mathrm{polylog}\left(\frac{1}{\delta},\ \frac{1}{\sigma_\eta^4},\ \tau,\ n_{\widehat{x}},\ ||\Sigma_0||,\ ||K'_{\mathcal{I}^i_{\widehat{Q}}}||,\ ||\mathfrak{P}_\infty||\right) \qquad (83)
\end{aligned}
$$

Next, consider

$$||\mathbb{E}[\psi_{t+1}|x_t,u_t]-\psi_{t+1}|| \leq \left\| \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} (\sigma_w^2\mathbb{I}-\widehat{w}_t^i\widehat{w}_t^{i\intercal}) \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal \right\| + \left\| \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \widehat{w}_t^i(\widehat{A}_t^i\widehat{x}_t^i+\widehat{B}_t^i\widehat{u}_t^i)^\intercal \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal \right\|_F$$

(By triangle inequality and the identity $||\cdot|| \leq ||\cdot||_F$.)

$$= \left\|(\sigma_w^2\mathbb{I}-\widehat{w}_t^i\widehat{w}_t^{i\intercal})\right\| \left\| \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal \right\| \left\| \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \right\| + \left\|\widehat{w}_t^i(\widehat{A}_t^i\widehat{x}_t^i+\widehat{B}_t^i\widehat{u}_t^i)^\intercal\right\| \left\| \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^\intercal \right\|_F \left\| \begin{bmatrix} I \\ K_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix} \right\|_F$$

[By sub-multiplicativity.]

$$\leq (1+||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)(||\sigma_w^2\mathbb{I}-\widehat{w}_t^i\widehat{w}_t^{i\intercal}||+n_{\widehat{x}}||\widehat{w}_t^i(\widehat{A}_t^i\widehat{x}_t^i+\widehat{B}_t^i\widehat{u}_t^i)^\intercal||_F) \tag{84}$$

By Corollary 5.50 in Vershynin (2010), we have that w.p $1-\delta/T$

$$\begin{aligned} ||\sigma_w^2\mathbb{I}-\widehat{w}_t^i\widehat{w}_t^{i\intercal}|| &\leq c_9\sigma_w^2 \max\left(\sqrt{n_{\widehat{x}}+\log(2T/\delta)}, n_{\widehat{x}}+\log(2T/\delta)\right) \\ &= c_9\sigma_w^2(n_{\widehat{x}}+\log(2T/\delta)). \end{aligned} \tag{85}$$

Next, consider

$$\begin{aligned} ||\widehat{w}_t^i(\widehat{A}_t^i\widehat{x}_t^i+\widehat{B}_t^i\widehat{u}_t^i)^\intercal||_F &= \sqrt{\mathrm{Tr}\left((\widehat{A}_t^i\widehat{x}_t^i+\widehat{B}_t^i\widehat{u}_t^i)(\widehat{w}_t^i)^\intercal\widehat{w}_t^i(\widehat{A}_t^i\widehat{x}_t^i+\widehat{B}_t^i\widehat{u}_t^i)^\intercal\right)} \\ &= \sqrt{\mathrm{Tr}\left((\widehat{w}_t^i)^\intercal\widehat{w}_t^i(\widehat{A}_t^i\widehat{x}_t^i+\widehat{B}_t^i\widehat{u}_t^i)^\intercal(\widehat{A}_t^i\widehat{x}_t^i+\widehat{B}_t^i\widehat{u}_t^i)\right)} \quad \text{(By cyclic property of trace.)} \\ &= \left\|\widehat{w}_t^i\right\|_2 \left\|\widehat{A}_t^i\widehat{x}_t^i+\widehat{B}_t^i\widehat{u}_t^i\right\|_2. \end{aligned} \tag{86}$$

Observe that $\left\|\widehat{w}_t^i\right\|_2^2 = \frac{1}{2}[a \;\; b]\begin{bmatrix} 0 & \sigma_w^2 \\ \sigma_w^2 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}$, $\left\|\widehat{A}_t^i\widehat{x}_t^i+\widehat{B}_t^i\widehat{u}_t^i\right\|_2^2 = \frac{1}{2}[a \;\; b]\begin{bmatrix} 0 & \Sigma_{t+1}^{\widehat{x}} \\ \Sigma_{t+1}^{\widehat{x}} & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}$, where a, b are isotropic Gaussian random variables. Then, by Hanson-Wright inequality (Theorem F.1), we have that w.p $1-\delta/T$

$$\left\|\widehat{w}_t^i\right\|_2^2 \leq \min\left(\frac{\sigma_w^2\,\|\mathbb{I}\|_F}{\sqrt{c_{10}}}\sqrt{\log(2T/\delta)}, \frac{\sigma_w^2\,\|\mathbb{I}\|}{c_{10}}\log(2T/\delta)\right) \overset{(a)}{\leq} c_{10}\sigma_w^2\log(T/\delta),$$

$$\left\|\widehat{A}_t^i\widehat{x}_t^i+\widehat{B}_t^i\widehat{u}_t^i\right\|_2^2 \leq \min\left(\frac{\left\|\Sigma_{t+1}^{\widehat{x}}\right\|_F}{\sqrt{c_{11}}}\sqrt{\log(2T/\delta)}, \frac{\left\|\Sigma_{t+1}^{\widehat{x}}\right\|}{c_{11}}\log(2T/\delta)\right) \overset{(b)}{\leq} c_{11}\left\|\Sigma_{t+1}^{\widehat{x}}\right\|\log(T/\delta),$$

where (a), (b) are due to the identities $||\cdot|| \leq ||\cdot||_F$, $\min\left(\log(2T/\delta)/c, \sqrt{\log(2T/\delta)/c}\right) \leq \log(2T/\delta)/c$. Therefore, from (86) we have that w.p $1-\delta/T$,

$$\begin{aligned} &||\widehat{w}_t^i(\widehat{A}_t^i\widehat{x}_t^i+\widehat{B}_t^i\widehat{u}_t^i)^\intercal||_F \leq \sqrt{c_{10}\sigma_w^2\log(T/\delta)}\sqrt{c_{11}\left\|\Sigma_{t+1}^{\widehat{x}}\right\|\log(T/\delta)} \\ &= \sqrt{c_{10}c_{11}}\sigma_w\log(T/\delta)\sqrt{\left\|L^{t+1}\Sigma_i^{\widehat{x}}(0)(L^\intercal)^{t+1}+\mathfrak{P}_{t+1}\right\|} \\ &= \sqrt{c_{10}c_{11}}\sigma_w\log(T/\delta)\sqrt{\left\|L^{t+1}\Sigma_i^{\widehat{x}}(0)(L^\intercal)^{t+1}+L\mathfrak{P}_tL^\intercal+\sigma_w^2\mathbb{I}+\sigma_\eta^2\widehat{B}\widehat{B}^\intercal\right\|} \\ &\overset{(a)}{\leq} \sqrt{c_{10}c_{11}}\sigma_w\log(T/\delta)\sqrt{\tau^2\rho^{2(t+1)}\left\|\Sigma_i^{\widehat{x}}(0)\right\|+\|\mathfrak{P}_\infty\|+\sigma_w^2+\left\|\sigma_\eta^2\widehat{B}\widehat{B}^\intercal\right\|}, \end{aligned} \tag{87}$$

where $L = \widehat{A} + \widehat{B}\widehat{K}'$, and (a) follows from the $(\tau, \rho)-$stability assumption of L, $\mathfrak{P}_t \preceq \mathfrak{P}_\infty$, and $L\mathfrak{P}_\infty L^\intercal \preceq \mathfrak{P}_\infty$.

Combining (84), (85), and (87) yields

$$
\begin{aligned}
&||\mathbb{E}[\psi_{t+1}|x_t, u_t] - \psi_{t+1}|| \leq \\
&(1 + ||K_{\mathcal{I}^i_{\widehat{Q}}}||^2)\left(c_9\sigma_w^2(n_{\widehat{x}} + \log(2T/\delta)) + n_{\widehat{x}}\sqrt{c_{10}c_{11}}\sigma_w \log(T/\delta)\sqrt{\tau^2\rho^{2(t+1)}\left\|\Sigma_i^{\widehat{x}}(0)\right\| + \|\mathfrak{P}_\infty\| + \sigma_w^2 + \left\|\sigma_\eta^2\widehat{B}\widehat{B}^\intercal\right\|}\right) \\
&\overset{(a)}{\leq} 2\sqrt{c_{10}c_{11}}\sigma_w(n_{\widehat{x}})\log(T/\delta)\sqrt{\tau^2\rho^{2(t+1)}\left\|\Sigma_i^{\widehat{x}}(0)\right\| + \|\mathfrak{P}_\infty\| + \sigma_w^2 + \left\|\sigma_\eta^2\widehat{B}\widehat{B}^\intercal\right\|}, \qquad (88)
\end{aligned}
$$

where (a) follows from $\sigma_w^2\mathbb{I} \leq \mathfrak{P}_\infty$ which implies $\sigma_w^2 \leq \|\mathfrak{P}_\infty\|^{\frac{1}{2}}$.

F.3. Obtaining the final bound

Next, observe that we can lower bound $\sigma_{\min}(\mathbb{I} - L \otimes_s L)$ using $(\tau, \rho)-$stability of $\widehat{A} + \widehat{B}\widehat{K}$ because for $k \geq 1$, we have

$$
\begin{aligned}
||L^k|| &= \left\|\begin{bmatrix} A_{\mathcal{I}^i_{\widehat{Q}}} & B_{\mathcal{I}^i_{\widehat{Q}}} \\ K_{\mathcal{I}^i_{\widehat{Q}}}A_{\mathcal{I}^i_{\widehat{Q}}} & K_{\mathcal{I}^i_{\widehat{Q}}}B_{\mathcal{I}^i_{\widehat{Q}}} \end{bmatrix}^k\right\| = \left\|\begin{bmatrix}\mathbb{I} \\ \widehat{K}\end{bmatrix}(\widehat{A} + \widehat{B}\widehat{K})^{k-1}\begin{bmatrix}\widehat{A} & \widehat{B}\end{bmatrix}\right\| \\
&\leq \left\|\begin{bmatrix}\mathbb{I} \\ \widehat{K}\end{bmatrix}\right\|\left\|(\widehat{A} + \widehat{B}\widehat{K})^{k-1}\begin{bmatrix}\widehat{A} & \widehat{B}\end{bmatrix}\right\| \text{ (Sub-multiplicativity)} \\
&\overset{(a)}{\leq} 2\tau\rho^{k-1}\max(1, ||\widehat{K}||)\left\|\begin{bmatrix}\widehat{A} & \widehat{B}\end{bmatrix}\right\| \\
&\leq \frac{2\tau\rho^k||\widehat{K}||_+\sqrt{(||\widehat{A}||^2 + ||\widehat{B}||^2)}}{\rho} \text{ (See Appendix H)}, \qquad (89)
\end{aligned}
$$

where $||\cdot||_+ = \max(1, ||\cdot||)$. The inequality (a) above follows from the identity that

$$
\begin{aligned}
\left\|\begin{bmatrix}\mathbb{I} \\ \widehat{K}\end{bmatrix}\right\| &= \sqrt{\lambda_{\max}(\mathbb{I} + \widehat{K}^\intercal\widehat{K})} = \sqrt{\max_{x \in S_{n_{\widehat{x}}}}\frac{x^\intercal(\mathbb{I} + \widehat{K}^\intercal\widehat{K})x}{x^\intercal x}} = \sqrt{1 + \max_{x \in S_{n_{\widehat{x}}}}\frac{x^\intercal\widehat{K}^\intercal\widehat{K}x}{x^\intercal x}} \\
&= \sqrt{1 + \lambda_{\max}(\widehat{K}^\intercal\widehat{K})} \leq 2\max\left(1, \sqrt{\lambda_{\max}(\widehat{K}^\intercal\widehat{K})}\right) = 2\max(1, ||\widehat{K}||). \qquad (90)
\end{aligned}
$$

Therefore, from (89) we conclude that L is $(\frac{2\tau\rho^k||\widehat{K}||_+\sqrt{(||\widehat{A}||^2+||\widehat{B}||^2)}}{\rho}, \rho)-$stable. Next, we know by definition of singular value that $\sigma_{\min}(\mathbb{I} - L \otimes_s L) = \frac{1}{||(\mathbb{I}-L\otimes_s L)^{-1}||}$. Therefore, for any unit norm vector v,

$$||(\mathbb{I}-L\otimes_s L)^{-1}\text{svec}(\text{smat}(v))|| = ||(\sum_{k=0}^{\infty}(L\otimes_s L)^k\text{svec}(\text{smat}(v))|| \quad \left(\because (\mathbb{I}-M)^{-1} = \sum_{k=0}^{\infty} M^k,\ \text{See Section 3.4, Peter}\right.$$

$$= \left\| \sum_{k=0}^{\infty} L^k \ \text{smat}(v)(L^\intercal)^k \right\| \ (\text{By definition of } \otimes_s)$$

$$\leq \sum_{k=0}^{\infty} \left\|L^k\right\| \|\ \text{smat}(v)\|_F \left\|(L^\intercal)^k\right\| \ \ (\text{By triangle inequality, sub-multiplicativity of } ||\cdot||, \text{ and } ||\cdot|| \leq ||\cdot||_F)$$

$$\leq \frac{4\tau^2||\widehat{K}||_+^2(||\widehat{A}||^2+||\widehat{B}||^2)}{\rho^2}\sum_{k=0}^{\infty}\rho^{2k}$$

$$= \frac{4\tau^2||\widehat{K}||_+^2(||\widehat{A}||^2+||\widehat{B}||^2)}{\rho^2(1-\rho^2)}. \tag{91}$$

Therefore, we have that

$$\sigma_{\min}(\mathbb{I}-L\otimes_s L) \geq \frac{\rho^2(1-\rho^2)}{4\tau^2||\widehat{K}||_+^2(||\widehat{A}||^2+||\widehat{B}||^2)}. \tag{92}$$

From (62), (65), we have that as long as $T \geq \tilde{O}(1)(n_{\widehat{x}}+n_{\widehat{u}})^2$

$$\sigma_{\min}(\mathbf{\Phi}) \geq c\frac{\sigma_\eta^2\sqrt{T}}{||\widehat{K}'||_+^2}. \tag{93}$$

By combining (83), (92) , (93) and the condition in (39) we obtain that

$$\frac{||(\mathbf{\Phi}^\intercal\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\intercal(\mathbf{\Xi}-\mathbf{\Psi}_+)||}{\sigma_{\min}(\mathbf{\Phi})\sigma_{\min}(\mathbb{I}-L\otimes_s L)} < \frac{1}{2}$$

$$\Rightarrow \frac{\tilde{O}(1)n_{\widehat{x}}(n_{\widehat{x}}+n_{\widehat{u}})||\widehat{K}'||_+^2}{\sigma_\eta^2\sqrt{T}}\sigma_w\sqrt{\tau^2\rho^4||\Sigma^{\widehat{x}}(0)||+||\mathfrak{P}_\infty||+\sigma_w^2+\sigma_\eta^2||\widehat{B}||^2}||\widehat{K}||_+^2\frac{4\tau^2||\widehat{K}||_+^2(||\widehat{A}||^2+||\widehat{B}||^2)}{\rho^2(1-\rho^2)} \leq \frac{1}{2}$$

$$\Rightarrow T \geq \frac{\tilde{O}(1)n_{\widehat{x}}^2(n_{\widehat{x}}+n_{\widehat{u}})^2||\widehat{K}'||_+^4}{\sigma_\eta^4}\sigma_w^2(\tau^2\rho^4||\Sigma^{\widehat{x}}(0)||+||\mathfrak{P}_\infty||+\sigma_w^2+\sigma_\eta^2||\widehat{B}||^2)\frac{\tau^4||\widehat{K}||_+^8(||\widehat{A}||^2+||\widehat{B}||^2)^2}{\rho^4(1-\rho^2)^2} \tag{94}$$

If T satisfies (94), then from (39) we have that

$$||p-\hat{p}|| \leq 2\frac{||(\mathbf{\Phi}^\intercal\mathbf{\Phi})^{-\frac{1}{2}}\mathbf{\Phi}^\intercal(\mathbf{\Xi}-\mathbf{\Psi}_+)p||}{\sigma_{\min}(\mathbf{\Phi})\sigma_{\min}(\mathbb{I}-L\otimes_s L)}$$

$$||p-\hat{p}|| \leq \frac{\tilde{O}(1)(n_{\widehat{x}}+n_{\widehat{u}})||\widehat{K}'||_+^2}{\sigma_\eta^2\sqrt{T}}\sigma_w\sqrt{\tau^2\rho^4||\Sigma^{\widehat{x}}(0)||+||\mathfrak{P}_\infty||+\sigma_w^2+\sigma_\eta^2||B||^2}||P_i||_F\frac{\tau^2||\widehat{K}||_+^4(||\widehat{A}||^2+||\widehat{B}||^2)}{\rho^2(1-\rho^2)}. \tag{95}$$

For sample complexity analysis, we are usually interested in finding the sufficient samples required to achieve $\epsilon-$optimal critic estimate i.e., $||p-\hat{p}|| \leq \epsilon$.

Set $\epsilon = \dfrac{\tilde{O}(1)W(n_{\widehat{x}} + n^i_{\widehat{u}})}{\sigma^2_\eta \sqrt{T}} ||P_i||_F$, which implies $T = \dfrac{(\tilde{O}(1))^2 W^2 (n_{\widehat{x}} + n^i_{\widehat{u}})^2}{\sigma^4_\eta \epsilon^2} ||P_i||^2_F \leq \dfrac{(\tilde{O}(1))^2 W^2 (n_{\widehat{x}} + n^i_{\widehat{u}})^3}{\sigma^4_\eta \epsilon^2} ||P$

Observe that

$$\tilde{O}(1)(n_{\widehat{x}} + n^i_{\widehat{u}})^2 \leq \frac{(\tilde{O}(1))^2 W^2 (n_{\widehat{x}} + n^i_{\widehat{u}})^3}{\sigma^4_\eta \epsilon^2} ||P_i||^2, \tag{96}$$

but $\dfrac{(\tilde{O}(1))^2 W^2 (n_{\widehat{x}} + n^i_{\widehat{u}})^3}{\sigma^4_\eta \epsilon^2} ||P_i||^2$ may not be greater than $\dfrac{\tilde{O}(1) W^2 (n_{\widehat{x}} + n^i_{\widehat{u}})^4}{\sigma^4_\eta}$.

Therefore, we can conclude that to achieve $||q_i - \hat{q}_i|| \leq \epsilon$, we require only

$$T \leq \max\left(\frac{(\tilde{O}(1))^2 W^2 (n_{\widehat{x}} + n^i_{\widehat{u}})^3}{\sigma^4_\eta \epsilon^2} ||P_i||^2, \frac{\tilde{O}(1) W^2 (n_{\widehat{x}} + n^i_{\widehat{u}})^4}{\sigma^4_\eta} \right) \text{ samples.} \tag{97}$$

Appendix G. Analysis of the indirect case

Define $n^i_x = n_x|\mathcal{I}^i_Q|$, and $n^i_u = n_u|\mathcal{I}^i_Q|$.

Corollary 1 *Consider $\delta \in (0,1)$. Let the initial global state and the global control (during sample generation) $\forall\, t$ satisfy $x(0) \sim \mathcal{N}(x_0, \Sigma_0)$, $u(t) = K^{play}x(t) + \eta_t$, $\eta(t) \sim \mathcal{N}(\mathbf{0}, \sigma^2_\eta \mathbb{I})$, and $\sigma_\eta \leq \sigma_w$. For each $i \in \mathcal{V}$, let $K^{play}_{\mathcal{I}^i_Q}$, $K_{\mathcal{I}^i_Q}$ stabilize $(A_{\mathcal{I}^i_Q}, B_{\mathcal{I}^i_Q})$. Assume that $A_{\mathcal{I}^i_Q} + B_{\mathcal{I}^i_Q} K_{\mathcal{I}^i_Q}$ and $A_{\mathcal{I}^i_Q} + B_{\mathcal{I}^i_Q} K^{play}_{\mathcal{I}^i_Q}$ are (τ, ρ)-stable. Let $\mathfrak{P}_\infty = \mathcal{L}\left(A_{\mathcal{I}^i_Q} + B_{\mathcal{I}^i_Q} K_{\mathcal{I}^i_Q}, \sigma^2_w \mathbb{I} + \sigma^2_\eta B_{\mathcal{I}^i_Q} B^\intercal_{\mathcal{I}^i_Q} \right)$ and $\bar{\sigma}_i = \sqrt{\tau^2 \rho^4 ||\Sigma^x(0)|| + ||\mathfrak{P}_\infty|| + \sigma^2_w + \sigma^2_\eta ||B_{\mathcal{I}^i_Q}||^2}$. Further, $\forall\, i \in \mathcal{V}$, let T_i denote the minimum number of samples required during learning. Suppose that*

$$T_i \geq \tilde{O}(1) \; max \left\{ (n^i_x + n^i_u)^2, \frac{(n^i_x)^2 (n^i_x + n^i_u)^2 ||K^{play}_{\mathcal{I}^j_Q}||^4_+}{\sigma^4_\eta} \sigma^2_w \bar{\sigma}^2_i \frac{\tau^4_i ||K_{\mathcal{I}^j_Q}||^8_+ (||A_{\mathcal{I}^j_Q}||^2 + ||B_{\mathcal{I}^j_Q}||^2)^2}{\rho^4_i (1 - \rho^2_i)^2} \right\}.$$

Then, with probability $1 - \delta$,

$$\left\| \hat{q}^{true}_i - \hat{q}^{indirect}_i \right\| \leq \sum_{j \in \mathcal{I}^i_{GD}} \frac{\tilde{O}(1)(n^j_x + n^j_u) ||K^{play}_{\mathcal{I}^j_Q}||^2_+}{\sigma^2_\eta \sqrt{T}} \sigma_w \bar{\sigma}_j ||Q^{true}_j||_F \frac{\tau^2_j ||K_{\mathcal{I}^j_Q}||^2_+ (||A_{\mathcal{I}^j_Q}||^2 + ||B_{\mathcal{I}^j_Q}||^2)}{\rho^2_j (1 - \rho^2_j)}$$

whenever $T \geq max\, \{T_j\}_{j \in \mathcal{I}^i_{GD}}$, where $\tilde{O}(1)$ hides polylog $\left(\dfrac{T}{\delta},\ \dfrac{1}{\sigma^4_\eta},\ \tau,\ n_x,\ ||\Sigma_0||,\ ||K^{play}_{\mathcal{I}^i_Q}||,\ ||\mathfrak{P}_\infty|| \right)$.

Proof Employing a *linear architecture*, the Q-function for each agent i can be expressed as

$$c_i(x_{\mathcal{I}^i_Q}(t), u_{\mathcal{I}^i_Q}(t)) =$$
$$\lambda + \left[\text{svec}\left(\begin{bmatrix} x_{\mathcal{I}^i_Q}(t) \\ u_{\mathcal{I}^i_Q}(t) \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}^i_Q}(t) \\ u_{\mathcal{I}^i_Q}(t) \end{bmatrix}^\intercal \right) - \mathbb{E}\left[\text{svec}\left(\begin{bmatrix} x_{\mathcal{I}^i_Q}(t{+}1) \\ u_{\mathcal{I}^i_Q}(t{+}1) \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}^i_Q}(t{+}1) \\ u_{\mathcal{I}^i_Q}(t{+}1) \end{bmatrix}^\intercal \right) \right] \right]^\intercal \text{svec}(Q_i), \tag{98}$$

where $\lambda \in \mathbb{R}$ is a free parameter to satisfy the fixed point equation. Let $\lambda = \left\langle Q_i, \begin{bmatrix} \sigma_w^2 \mathbb{I}_{n_x|I_Q^i|} & \sigma_w^2 K_{\mathcal{I}_Q^i}^\intercal \\ \sigma_w^2 K_{\mathcal{I}_Q^i} & \sigma_w^2 K_{\mathcal{I}_{\widehat{Q}}^i} K_{\mathcal{I}_Q^i}^\intercal \end{bmatrix} \right\rangle$.

For brevity, in the remainder of the proof denote $\phi_t = \text{svec}\left(\begin{bmatrix} x_{\mathcal{I}_Q^i}(t) \\ u_{\mathcal{I}_Q^i}(t) \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}_Q^i}(t) \\ u_{\mathcal{I}_Q^i}(t) \end{bmatrix}^\intercal \right)$,

$\psi_t = \text{svec}\left(\begin{bmatrix} x_{\mathcal{I}_Q^i}(t) \\ K_{\mathcal{I}_Q^i} x_{\mathcal{I}_Q^i}(t) \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}_Q^i}(t) \\ K_{\mathcal{I}_Q^i} x_{\mathcal{I}_Q^i}(t) \end{bmatrix}^\intercal \right)$, $f = \text{svec}\left(\begin{bmatrix} \Sigma_{\mathcal{I}_Q^i}^x & \Sigma_{\mathcal{I}_Q^i}^x K_{\mathcal{I}_Q^i}^\intercal \\ K_{\mathcal{I}_Q^i} \Sigma_{\mathcal{I}_Q^i}^x & K_{\mathcal{I}_{\widehat{Q}}^i} \Sigma_{\mathcal{I}_Q^i}^x K_{\mathcal{I}_Q^i}^\intercal \end{bmatrix} \right)$, and $\xi_t = \mathbb{E}\left[\text{svec}\left(\begin{bmatrix} x_{\mathcal{I}_Q^i}(t+1) \\ u_{\mathcal{I}_Q^i}(t+1) \end{bmatrix} \begin{bmatrix} x_{\mathcal{I}_Q^i}(t+1) \\ u_{\mathcal{I}_Q^i}(t+1) \end{bmatrix}^\intercal \right) \right]$.

For a single trajectory $\left\{ x_{\mathcal{I}_Q^i}(t), u_{\mathcal{I}_Q^i}(t), x_{\mathcal{I}_Q^i}(t+1) \right\}_{t=1}^{T_i}$, the bellman equation for agent i can be expressed in matrix form as

$$\mathbf{c} = (\mathbf{\Phi} - \mathbf{\Xi} + \mathbf{F}) q_i, \tag{99}$$

$$\text{where } \mathbf{\Phi} = \begin{bmatrix} \phi_1^\intercal \\ \phi_2^\intercal \\ \vdots \\ \phi_{T_i}^\intercal \end{bmatrix}, \mathbf{\Xi} = \begin{bmatrix} \xi_1^\intercal \\ \xi_2^\intercal \\ \vdots \\ \xi_{T_i}^\intercal \end{bmatrix}, \mathbf{c} = \begin{bmatrix} c_i(1) \\ c_i(2) \\ \vdots \\ c_i(T_i) \end{bmatrix}, \mathbf{F} = \begin{bmatrix} f^\intercal \\ f^\intercal \\ \vdots \\ f^\intercal \end{bmatrix}$$

Observe that (99) is analogous to (30). Thus, we can conclude that if T_i satisfies

$$T_i \geq \tilde{O}(1) \max \left\{ (n_x^i + n_u^i)^2, \frac{(n_x^i)^2 (n_x^i + n_u^i)^2 ||K_{\mathcal{I}_Q^j}^{\text{play}}||_+^4}{\sigma_\eta^4} \sigma_w^2 \bar{\sigma}_i^2 \frac{\tau_i^4 ||K_{\mathcal{I}_Q^j}||_+^8 (||A_{\mathcal{I}_Q^j}||^2 + ||B_{\mathcal{I}_Q^j}||^2)^2}{\rho_i^4 (1-\rho_i^2)^2} \right\}, \tag{100}$$

where $\bar{\sigma}_i = \sqrt{\tau^2 \rho^4 ||\Sigma^x(0)|| + ||\mathfrak{P}_\infty|| + \sigma_w^2 + \sigma_\eta^2 ||B_{\mathcal{I}_Q^i}||^2}$. Then,

$$||q_i - \hat{q}_i|| \leq \frac{\tilde{O}(1)(n_x^i + n_u^i)||K_Q^{i\prime}||_+^2}{\sigma_\eta^2 \sqrt{T_i}} \sigma_w \bar{\sigma}_i ||K_Q^i||_+^2 ||Q_i||_F \frac{\tau_i^2 ||K_Q^i||_+^4 (||A_Q^i||^2 + ||B_Q^i||^2)}{\rho_i^2 (1-\rho_i^2)} \tag{101}$$

Define an operator M =ssmat(m, S, n) that maps $m \in \mathbb{R}^d$ to the corresponding entries M_{ij} $\forall \{(i,j) \in S \times S | i \leq j\}$. It is straightforward to verify that $\|m\| = \|\text{svec}(\text{ssmat}(m, \cdot, \cdot))\|$. Similar to Syed and Bai (2025), let Y denote the transformation that maps a vector to its non-zero subset and Y' be its inverse transformation. Hence, $Y'(m) = \text{svec}(\text{ssmat}(m, \cdot, \cdot))$, $Y(\text{svec}(\text{ssmat}(m, \cdot, \cdot))) = m$.

Then, the error in estimation of decomposed $\hat{Q}_i$ function can be expressed as

$$\begin{aligned}
\left\| q_{\widehat{Q}}^{\mathrm{dec},i} - \hat{q}_{\widehat{Q}}^{\mathrm{dec},i} \right\| &= \left\| Y \left(\operatorname{svec} \left(\sum_{j \in \mathcal{I}_{\mathrm{GD}}^i} \operatorname{ssmat}(q_j, \mathcal{I}_Q^j, \widehat{n}_x + \widehat{n}_u) \right) \right) - Y \left(\operatorname{svec} \left(\sum_{j \in \mathcal{I}_{\mathrm{GD}}^i} \operatorname{ssmat}(\hat{q}_j, \mathcal{I}_Q^j, \widehat{n}_x + \widehat{n}_u) \right) \right) \right\| \\
&= \left\| \sum_{j \in \mathcal{I}_{\mathrm{GD}}^i} Y \left(\operatorname{svec} \left(\operatorname{ssmat}(q_j, \mathcal{I}_Q^j, \widehat{n}_x + \widehat{n}_u) \right) - \operatorname{svec} \left(\operatorname{ssmat}(\hat{q}_j, \mathcal{I}_Q^j, \widehat{n}_x + \widehat{n}_u) \right) \right) \right\| \\
&\leq \sum_{j \in \mathcal{I}_{\mathrm{GD}}^i} \| q_j - \hat{q}_j \| \quad \text{(Using triangle inequality, and } \|m\| = \|\operatorname{svec}(\operatorname{ssmat}(m, \cdot, \cdot))\|\text{)}.
\end{aligned} \tag{102}$$

Combining (100), (101), and (102), we obtain that whenever the length of trajectory (number of samples) satisfies

$$T \geq \max \{T_j\}_{j \in \mathcal{I}_{\mathrm{GD}}^i} \tag{103}$$

then

$$\left\| q_{\widehat{Q}}^{\mathrm{dec},i} - \hat{q}_{\widehat{Q}}^{\mathrm{dec},i} \right\| \leq \sum_{j \in \mathcal{I}_{\mathrm{GD}}^i} \frac{\tilde{O}(1)(n_x^j + n_u^j)||K_Q^{j\prime}||_+^2}{\sigma_\eta^2 \sqrt{T_j}} \sigma_w \sqrt{\tau_j^2 \rho_j^4 |||\Sigma_Q^{x,j}(0)|| + ||\mathfrak{P}_\infty|| + \sigma_w^2 + \sigma_\eta^2 ||B_Q^j||^2} ||K_Q^j||_+^2 ||Q_j||_F \frac{\tau_j^2 ||K_Q^j||_+^2 (||A_Q^j||^2 +}{\rho_j^2 (1 - \rho_j^2)} \tag{104}$$

Let $g_i = |\mathcal{I}_{\mathrm{GD}}^i|$. Then, define $w_1,\ w_2,\ \cdots,\ w_{g_i} \in [0, 1]$ such that $\sum_k^{g_i} w_k = 1$. Then, from (**??**), we conclude that to achieve $\left\| q_{\widehat{Q}}^{\mathrm{dec},i} - \hat{q}_{\widehat{Q}}^{\mathrm{dec},i} \right\| \leq \epsilon$, we require only

$$T \leq \max_{j \in \mathcal{I}_{\mathrm{GD}}^i} \left(\max \left(\frac{(\tilde{O}(1))^2 W_j^2 (n_x^j + n_u^j)^3}{\sigma_\eta^4 w_j^2 \epsilon^2} ||Q_j||^2, \frac{\tilde{O}(1) W_j^2 (n_x^j + n_u^j)^4}{\sigma_\eta^4} \right) \right) \text{ samples.} \tag{105}$$

Therefore, from (97), (105), we can conclude that the decomposition of the Q-function is sample efficient provided that $\max\limits_{j \in \mathcal{I}_{\mathrm{GD}}^i} (n_x^j + n_u^j) < (n_{\widehat{x}} + n_{\widehat{u}})$ and the worst case sample complexity of the decomposition is equal to the undecomposed case.

Appendix H. Bound on spectral norm of a block matrix

https://math.stackexchange.com/questions/2006773/spectral-norm-of-concatenation-

Define $M = \begin{bmatrix} A^\intercal A & A^\intercal B \\ B^\intercal A & B^\intercal B \end{bmatrix}$. Then, observe that $\left\| \begin{bmatrix} \widehat{A} & \widehat{B} \end{bmatrix} \right\| = \sqrt{\lambda_{\max}(M)}$. Since M is symmetric, by Courant-Fischer theorem, we have that

$$\lambda_{\max}(M) = \max_{x : ||x|| = 1} \frac{x^\intercal M x}{x^\intercal x}.$$

Let $x = [u^\intercal \;\; v^\intercal]$, where $u,\ v$ are eigenvectors of $A^\intercal A,\ B^\intercal B$ respectively. **Note that** $\|u\|^2 + \|v\|^2 = 1$**, so you should be able to tighten the bound.** Then,

$$\begin{aligned}
x^\intercal M x &= u^\intercal A^\intercal A u + v^\intercal B^\intercal A u + u^\intercal A^\intercal B v + v^\intercal B^\intercal B v \\
&\leq u^\intercal A^\intercal A u + 2||Au||\,||Bv|| + v^\intercal B^\intercal B v \text{ (By Cauchy-Schwarz inequality)} \\
&\leq 2\left\|Au\right\|^2 + 2\left\|Bv\right\|^2 \text{ (Since, } 2||Au||\,||Bv|| \leq ||Au||^2 + ||Bv||^2. \text{)} \\
&\leq 2||A||^2||u||^2 + 2||B||^2||v||^2 \\
&\leq 2(||A||^2 + ||B||^2)(||u||^2 + ||v||^2) \;(\because ||A||^2 + ||B||^2 \geq \{||A||^2, ||B||^2\}.) \\
x^\intercal M x &= ||A||^2 + ||B||^2 \;(\because ||u||^2 + ||v||^2 = 1.) \\
\Rightarrow \lambda_{\text{max}}(M) &= ||A||^2 + ||B||^2 \\
\Rightarrow \left\|\begin{bmatrix}\widehat{A} & \widehat{B}\end{bmatrix}\right\| &\leq \sqrt{(||A||^2 + ||B||^2)}.
\end{aligned} \tag{106}$$

■