Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

# Supplementary materials for

Zhenyi ZHANG, Jie HUANG, Congjie PAN, 2024. Multi-agent reinforcement learning behavioral control for nonlinear second-order systems. *Front Inform Technol Electron Eng*, 25(6):869-886.
https://doi.org/10.1631/FITEE.2300394

## 1 Proof of Theorem 1

Consider the Lyapunov function candidate of Theorem 1 as

$$V_u = \frac{1}{2}\sum_{i=1}^{N} \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{\Upsilon}_{\mathrm{e},i} \boldsymbol{e}_i + \frac{1}{2}\sum_{i=1}^{N} \boldsymbol{\tau}_i^{\mathrm{T}} \boldsymbol{\tau}_i + \frac{1}{2}\sum_{i=1}^{N} \mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}^{\mathrm{T}} \boldsymbol{\Upsilon}_{\mathrm{f},i}^{-1} \tilde{\boldsymbol{\omega}}_{\mathrm{f},i}\right\} + \frac{1}{2}\sum_{i=1}^{N} \mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}} \tilde{\boldsymbol{\omega}}_{\mathrm{c},i}\right\} + \frac{1}{2}\sum_{i=1}^{N} \mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}} \tilde{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}, \tag{S1}$$

where $\boldsymbol{\Upsilon}_{\mathrm{e},i} = \begin{bmatrix} \xi_{\mathrm{e},i}\boldsymbol{I}_n & \boldsymbol{I}_n \\ \boldsymbol{I}_n & \boldsymbol{I}_n \end{bmatrix}$ is a positive-definite matrix if condition (42) of Theorem 1 is held; $\xi_{\mathrm{e},i}$ is a designed parameter of $\boldsymbol{\Upsilon}_{\mathrm{e},i}$; $\mathrm{tr}\{\cdot\}$ is the trace of a matrix; and $\tilde{\boldsymbol{\omega}}_{\mathrm{f},i} = \hat{\boldsymbol{\omega}}_{\mathrm{f},i} - \boldsymbol{\omega}_{\mathrm{f},i}^{*}$, $\tilde{\boldsymbol{\omega}}_{\mathrm{c},i} = \hat{\boldsymbol{\omega}}_{\mathrm{c},i} - \boldsymbol{\omega}_{\mathrm{V},i}^{*}$, and $\tilde{\boldsymbol{\omega}}_{\mathrm{a},i} = \hat{\boldsymbol{\omega}}_{\mathrm{a},i} - \boldsymbol{\omega}_{\mathrm{V},i}^{*}$ are the weight errors of the identifier, critic, and actor, respectively.

The following result is yielded by differentiating Eq. (S1) along Eqs. (21), (30), (37), and (39)–(41):

$$\begin{aligned}
\dot{V}_u =& \sum_{i=1}^{N}\Big[\xi_{\mathrm{e},i}\boldsymbol{e}_{\mathrm{p},i}^{\mathrm{T}}\boldsymbol{e}_{\mathrm{v},i} + \boldsymbol{e}_{\mathrm{v},i}^{\mathrm{T}}\boldsymbol{e}_{\mathrm{v},i} + (\boldsymbol{e}_{\mathrm{p},i}^{\mathrm{T}} + \boldsymbol{e}_{\mathrm{v},i}^{\mathrm{T}})(-\xi_{\mathrm{p},i}\boldsymbol{e}_{\mathrm{p},i} - \xi_{\mathrm{v},i}\boldsymbol{e}_{\mathrm{v},i} - \hat{\boldsymbol{\omega}}_{\mathrm{f},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{f},i}(\boldsymbol{\chi}_i) - \boldsymbol{\tau}_i - \frac{1}{2\beta_{\mathrm{V}}}\hat{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i) \\
&+ \boldsymbol{f}_i(\boldsymbol{\chi}_i) - \dot{\boldsymbol{v}}_{i,\mathrm{r}})\Big] + \sum_{i=1}^{N}\boldsymbol{\tau}_i^{\mathrm{T}}(-\xi_{\tau,i}\boldsymbol{\tau}_i + \boldsymbol{u}_{\Delta,i}) + \sum_{i=1}^{N}\mathrm{tr}\Big\{\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}^{\mathrm{T}}\big[\boldsymbol{\Psi}_{\mathrm{f},i}(\boldsymbol{\chi}_i)(\boldsymbol{e}_{\mathrm{p},i}^{\mathrm{T}} + \boldsymbol{e}_{\mathrm{v},i}^{\mathrm{T}}) - \boldsymbol{\nu}_{\mathrm{f},i}\hat{\boldsymbol{\omega}}_{\mathrm{f},i}\big]\Big\} \\
&- \sum_{i=1}^{N}\mathrm{tr}\Big\{\gamma_{\mathrm{c},i}\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{c},i}\Big\} \\
&- \sum_{i=1}^{N}\mathrm{tr}\Big\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\big[\gamma_{\mathrm{a},i}(\hat{\boldsymbol{\omega}}_{\mathrm{a},i} - \hat{\boldsymbol{\omega}}_{\mathrm{c},i}) + \gamma_{\mathrm{c},i}\hat{\boldsymbol{\omega}}_{\mathrm{c},i}\big]\Big\}.
\end{aligned} \tag{S2}$$

According to Property 1 and Eq. (32), Eq. (S2) can be further simplified as

$$\begin{aligned}
\dot{V}_u =& -\sum_{i=1}^{N}\left[\xi_{\mathrm{p},i}\left\|\boldsymbol{e}_{\mathrm{p}}\right\|^2 + (\xi_{\mathrm{v},i}-1)\left\|\boldsymbol{e}_{\mathrm{v}}\right\|^2 + \xi_{\tau,i}\left\|\boldsymbol{\tau}_i\right\|^2\right] - \sum_{i=1}^{N}(\xi_{\mathrm{p},i} + \xi_{\mathrm{v},i} - \xi_{\mathrm{e},i})\boldsymbol{e}_{\mathrm{p},i}^{\mathrm{T}}\boldsymbol{e}_{\mathrm{v},i} - \sum_{i=1}^{N}(\boldsymbol{e}_{\mathrm{p},i}^{\mathrm{T}} + \boldsymbol{e}_{\mathrm{v},i}^{\mathrm{T}}) \\
&\cdot\left[\boldsymbol{\tau}_i + \frac{1}{2\beta_{\mathrm{V}}}\hat{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i) + \dot{\boldsymbol{v}}_{i,\mathrm{r}} - \boldsymbol{\sigma}_{\mathrm{f},i}(\boldsymbol{\chi}_i)\right] + \sum_{i=1}^{N}\boldsymbol{\tau}_i^{\mathrm{T}}\boldsymbol{u}_{\Delta,i} - \sum_{i=1}^{N}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}^{\mathrm{T}}\boldsymbol{\nu}_{\mathrm{f},i}\hat{\boldsymbol{\omega}}_{\mathrm{f},i}\right\} \\
&- \sum_{i=1}^{N}\gamma_{\mathrm{c},i}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi},\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{c},i}\right\} - \sum_{i=1}^{N}\gamma_{\mathrm{a},i}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{a},i}\right\} \\
&+ \sum_{i=1}^{N}(\gamma_{\mathrm{a},i} - \gamma_{\mathrm{c},i})\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{c},i}\right\}.
\end{aligned} \tag{S3}$$

The following results can be obtained by using Properties 1 and 2:

$$-\left(\boldsymbol{e}_{\mathrm{p},i}^{\mathrm{T}} + \boldsymbol{e}_{\mathrm{v},i}^{\mathrm{T}}\right)\boldsymbol{\tau}_i \le \frac{1}{2}\left\|\boldsymbol{e}_{\mathrm{p},i}\right\|^2 + \frac{1}{2}\left\|\boldsymbol{e}_{\mathrm{v},i}\right\|^2 + \left\|\boldsymbol{\tau}_i\right\|^2, \tag{S4}$$

$$-\frac{1}{2\beta_{\mathrm{V}}}(\boldsymbol{e}_{\mathrm{p},i}^{\mathrm{T}}+\boldsymbol{e}_{\mathrm{v},i}^{\mathrm{T}})\hat{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\leq\frac{1}{2}\left\|\boldsymbol{e}_{\mathrm{p},i}\right\|^2+\frac{1}{2}\left\|\boldsymbol{e}_{\mathrm{v},i}\right\|^2+\frac{1}{4\beta_{\mathrm{V}}^2}\mathrm{tr}\left\{\hat{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}, \tag{S5}$$

$$-(\boldsymbol{e}_{\mathrm{p},i}^{\mathrm{T}}+\boldsymbol{e}_{\mathrm{v},i}^{\mathrm{T}})\dot{\boldsymbol{v}}_{i,\mathrm{r}}\leq\frac{1}{2}\left\|\boldsymbol{e}_{\mathrm{p},i}\right\|^2+\frac{1}{2}\left\|\boldsymbol{e}_{\mathrm{v},i}\right\|^2+\left\|\dot{\boldsymbol{v}}_{i,\mathrm{r}}\right\|^2, \tag{S6}$$

$$(\boldsymbol{e}_{\mathrm{p},i}^{\mathrm{T}}+\boldsymbol{e}_{\mathrm{v},i}^{\mathrm{T}})\boldsymbol{\sigma}_{\mathrm{f},i}(\boldsymbol{\chi}_i)\leq\frac{1}{2}\left\|\boldsymbol{e}_{\mathrm{p},i}\right\|^2+\frac{1}{2}\left\|\boldsymbol{e}_{\mathrm{v},i}\right\|^2+\left\|\boldsymbol{\sigma}_{\mathrm{f},i}(\boldsymbol{\chi}_i)\right\|^2\leq\frac{1}{2}\left\|\boldsymbol{e}_{\mathrm{p},i}\right\|^2+\frac{1}{2}\left\|\boldsymbol{e}_{\mathrm{v},i}\right\|^2+\varepsilon_{\mathrm{f}}^2, \tag{S7}$$

$$\boldsymbol{\tau}_i^{\mathrm{T}}\boldsymbol{u}_{\Delta,i}\leq\frac{1}{2}\left\|\boldsymbol{\tau}_i\right\|^2+\frac{1}{2}\left\|\boldsymbol{u}_{\Delta,i}\right\|^2\leq\frac{1}{2}\left\|\boldsymbol{\tau}_i\right\|^2+\frac{1}{2}\varepsilon_{\Delta}^2. \tag{S8}$$

Substituting inequalities (S4)–(S8) into Eq. (S3), we obtain

$$\begin{aligned}\dot{V}_u\leq&-\sum_{i=1}^{N}\left[(\xi_{\mathrm{p},i}-2)\left\|\boldsymbol{e}_{\mathrm{p},i}\right\|^2+(\xi_{\mathrm{v},i}-3)\left\|\boldsymbol{e}_{\mathrm{v},i}\right\|^2+\left(\xi_{\tau,i}-\frac{3}{2}\right)\left\|\boldsymbol{\tau}_i\right\|^2\right]-\sum_{i=1}^{N}(\xi_{\mathrm{p},i}+\xi_{\mathrm{v},i}-\xi_{\mathrm{e},i})\boldsymbol{e}_{\mathrm{p},i}^{\mathrm{T}}\boldsymbol{e}_{\mathrm{v},i}\\&-\sum_{i=1}^{N}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}^{\mathrm{T}}\nu_{\mathrm{f},i}\hat{\boldsymbol{\omega}}_{\mathrm{f},i}\right\}-\sum_{i=1}^{N}\gamma_{\mathrm{c},i}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{c},i}\right\}-\sum_{i=1}^{N}\gamma_{\mathrm{a},i}\mathrm{tr}\big\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\\&\cdot\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{a},i}\big\}+\sum_{i=1}^{N}(\gamma_{\mathrm{a},i}-\gamma_{\mathrm{c},i})\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{c},i}\right\}\\&+\frac{1}{4\beta_{\mathrm{V}}^2}\sum_{i=1}^{N}\mathrm{tr}\left\{\hat{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi},\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}+\sum_{i=1}^{N}\left\|\dot{\boldsymbol{v}}_{i,\mathrm{r}}\right\|^2+N\varepsilon_{\mathrm{f}}^2+\frac{N}{2}\varepsilon_{\Delta}^2.\end{aligned} \tag{S9}$$

The results obtained pursuant to the use of Properties 1–3, and from the fact that $\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}=\hat{\boldsymbol{\omega}}_{\mathrm{f},i}-\boldsymbol{\omega}_{\mathrm{f},i}^*$, $\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}=\hat{\boldsymbol{\omega}}_{\mathrm{c},i}-\boldsymbol{\omega}_{\mathrm{V},i}^*$, and $\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}=\hat{\boldsymbol{\omega}}_{\mathrm{a},i}-\boldsymbol{\omega}_{\mathrm{V},i}^*$, are the following:

$$\mathrm{tr}\{\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}^{\mathrm{T}}\hat{\boldsymbol{\omega}}_{\mathrm{f},i}\}=\frac{1}{2}\mathrm{tr}\{\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}^{\mathrm{T}}\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}\}+\frac{1}{2}\mathrm{tr}\{\hat{\boldsymbol{\omega}}_{\mathrm{f},i}^{\mathrm{T}}\hat{\boldsymbol{\omega}}_{\mathrm{f},i}\}-\frac{1}{2}\mathrm{tr}\{\left(\boldsymbol{\omega}_{\mathrm{f},i}^*\right)^{\mathrm{T}}\boldsymbol{\omega}_{\mathrm{f},i}^*\}, \tag{S10}$$

$$\begin{aligned}&\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{c},i}\right\}=\frac{1}{2}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}\right\}\\&+\frac{1}{2}\mathrm{tr}\left\{\hat{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{c},i}\right\}-\frac{1}{2}\mathrm{tr}\left\{\left(\boldsymbol{\omega}_{\mathrm{V},i}^*\right)^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\omega}_{\mathrm{V},i}^*\right\},\end{aligned} \tag{S11}$$

$$\begin{aligned}&\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}=\frac{1}{2}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}\\&+\frac{1}{2}\mathrm{tr}\left\{\hat{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}-\frac{1}{2}\mathrm{tr}\left\{\left(\boldsymbol{\omega}_{\mathrm{V},i}^*\right)^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\omega}_{\mathrm{V},i}^*\right\},\end{aligned} \tag{S12}$$

$$\begin{aligned}&(\gamma_{\mathrm{a},i}-\gamma_{\mathrm{c},i})\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{c},i}\right\}\\&\leq\frac{\gamma_{\mathrm{a},i}-\gamma_{\mathrm{c},i}}{2}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}+\frac{\gamma_{\mathrm{a},i}-\gamma_{\mathrm{c},i}}{2}\mathrm{tr}\left\{\hat{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{c},i}\right\}.\end{aligned} \tag{S13}$$

Substituting Eqs. (S10)–(S12) into inequality (S9), we obtain

$$\begin{aligned}\dot{V}_u\leq&-\sum_{i=1}^{N}\frac{1}{2}\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{\Upsilon}_{\xi,i}\boldsymbol{e}_i-\sum_{i=1}^{N}\left(\xi_{\tau,i}-\frac{3}{2}\right)\left\|\boldsymbol{\tau}_i\right\|^2-\sum_{i=1}^{N}\frac{\nu_{\mathrm{f},i}}{2\overline{\lambda}_{\mathrm{f}}}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}^{\mathrm{T}}\boldsymbol{\Upsilon}_{\mathrm{f},i}^{-1}\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}\right\}\\&-\sum_{i=1}^{N}\frac{\gamma_{\mathrm{c},i}}{2}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi},\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}\right\}-\sum_{i=1}^{N}\frac{\gamma_{\mathrm{c},i}}{2}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi},\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}\\&-\sum_{i=1}^{N}\left(\gamma_{\mathrm{c},i}-\frac{\gamma_{\mathrm{a},i}}{2}\right)\mathrm{tr}\left\{\hat{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{c},i}\right\}\\&-\sum_{i=1}^{N}\left(\frac{\gamma_{\mathrm{a},i}}{2}-\frac{1}{4\beta_{\mathrm{V}}^2}\right)\mathrm{tr}\left\{\hat{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}+\Xi,\end{aligned} \tag{S14}$$

where

$$\boldsymbol{\Upsilon}_{\xi,i}=\begin{bmatrix}2(\xi_{\mathrm{p},i}-2)\boldsymbol{I}_n & (\xi_{\mathrm{p},i}+\xi_{\mathrm{v},i}-\xi_{\mathrm{e},i})\boldsymbol{I}_n\\ (\xi_{\mathrm{p},i}+\xi_{\mathrm{v},i}-\xi_{\mathrm{e},i})\boldsymbol{I}_n & 2(\xi_{\mathrm{v},i}-3)\boldsymbol{I}_n\end{bmatrix},$$

$$\begin{aligned}\Xi=&\sum_{i=1}^{N}\frac{\gamma_{\mathrm{a},i}+\gamma_{\mathrm{c},i}}{2}\mathrm{tr}\left\{\left(\boldsymbol{\omega}_{\mathrm{V},i}^{*}\right)^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\omega}_{\mathrm{V},i}^{*}\right\}\\&+\frac{1}{2}\sum_{i=1}^{N}\nu_{\mathrm{f},i}\mathrm{tr}\left\{\left(\boldsymbol{\omega}_{\mathrm{f},i}^{*}\right)^{\mathrm{T}}\boldsymbol{\omega}_{\mathrm{f},i}^{*}\right\}+\sum_{i=1}^{N}\left\|\dot{\boldsymbol{v}}_{i,\mathrm{r}}\right\|^2+N\varepsilon_{\mathrm{f}}^2+\frac{N}{2}\varepsilon_{\Delta}^2,\end{aligned}$$

since all terms are bounded, $\Xi$ is bounded as $|\Xi|\leq\vartheta_2$, and $\overline{\lambda}_{\mathrm{f}}$ is the maximal eigenvalue of $\boldsymbol{\Upsilon}_{\mathrm{f},i}^{-1}$.

Based on condition (42), inequality (S14) can be rewritten in the compact form as

$$\begin{aligned}\dot{V}_u\leq&-\frac{1}{2}\sum_{i=1}^{N}\frac{\underline{\lambda}_{\xi}}{\overline{\lambda}_{\mathrm{e}}}\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{\Upsilon}_{\mathrm{e},i}\boldsymbol{e}_i-\frac{1}{2}\sum_{i=1}^{N}\left[2\left(\xi_{\tau,i}-\frac{3}{2}\right)\right]\left\|\boldsymbol{\tau}_i\right\|^2-\frac{1}{2}\sum_{i=1}^{N}\frac{\nu_{\mathrm{f},i}}{\overline{\lambda}_{\mathrm{f}}}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}^{\mathrm{T}}\boldsymbol{\Upsilon}_{\mathrm{f},i}^{-1}\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}\right\}\\&-\frac{1}{2}\sum_{i=1}^{N}\underline{\lambda}_{\Psi}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}}\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}\right\}-\frac{1}{2}\sum_{i=1}^{N}\underline{\lambda}_{\Psi}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}+\vartheta_2,\end{aligned}\tag{S15}$$

where $\underline{\lambda}_{\xi}$ is the minimal eigenvalue of $\boldsymbol{\Upsilon}_{\xi,i}$, $\overline{\lambda}_{\mathrm{e}}$ is the maximal eigenvalue of $\boldsymbol{\Upsilon}_{\mathrm{e},i}$, and $\underline{\lambda}_{\Psi}$ is the minimal eigenvalue of $\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)$.

Let $\vartheta_1=\min\left\{\frac{\underline{\lambda}_{\xi}}{\overline{\lambda}_{e}},2\left(\xi_{\tau,i}-\frac{3}{2}\right),\frac{\nu_{\mathrm{f},i}}{\overline{\lambda}_{\mathrm{f}}},\underline{\lambda}_{\Psi}\right\}$. Then we have

$$\dot{V}_u\leq-\vartheta_1V_u+\vartheta_2.\tag{S16}$$

The final result is obtained by using Lemma 1 as

$$V_u\leq\mathrm{e}^{-\vartheta_1t}V_u(0)+\frac{\vartheta_2}{\vartheta_1}(1-\mathrm{e}^{-\vartheta_1t}).\tag{S17}$$

Therefore, Theorem 1 is proven completely.

## 2 Proof of mission stability

Let us consider the Lyapunov function candidate of mission stability as

$$V_\rho=\frac{1}{2}\tilde{\boldsymbol{\rho}}_{\mathrm{OA}}^{\mathrm{T}}\kappa_{\mathrm{OA}}\tilde{\boldsymbol{\rho}}_{\mathrm{OA}}+\frac{1}{2}\tilde{\boldsymbol{\rho}}_{\mathrm{FM}}^{\mathrm{T}}\kappa_{\mathrm{FM}}\tilde{\boldsymbol{\rho}}_{\mathrm{FM}}+\frac{1}{2}\tilde{\boldsymbol{\rho}}_{\mathrm{FR}}^{\mathrm{T}}\kappa_{\mathrm{FR}}\tilde{\boldsymbol{\rho}}_{\mathrm{FR}},\tag{S18}$$

where $\kappa_{\mathrm{OA}}$, $\kappa_{\mathrm{FM}}$, and $\kappa_{\mathrm{FR}}$ are the designed positive constants.

We assume that the behavior priority satisfies OA>FM>FR; then the differential of Eq. (S18) is

$$\dot{V}_\rho=-\frac{1}{2}\tilde{\boldsymbol{\zeta}}^{\mathrm{T}}\boldsymbol{\Gamma}_\rho\tilde{\boldsymbol{\zeta}},\tag{S19}$$

where

$$\boldsymbol{\Gamma}_\rho=\begin{bmatrix}\boldsymbol{\Gamma}_{\rho,11} & \boldsymbol{\Gamma}_{\rho,12} & \boldsymbol{\Gamma}_{\rho,13}\\ \boldsymbol{\Gamma}_{\rho,21} & \boldsymbol{\Gamma}_{\rho,22} & \boldsymbol{\Gamma}_{\rho,23}\\ \boldsymbol{\Gamma}_{\rho,31} & \boldsymbol{\Gamma}_{\rho,32} & \boldsymbol{\Gamma}_{\rho,33}\end{bmatrix}$$

with $\boldsymbol{\Gamma}_{\rho,11}=\kappa_{\mathrm{OA}}\boldsymbol{\Lambda}_{\mathrm{OA}}$, $\boldsymbol{\Gamma}_{\rho,21}=\boldsymbol{\Gamma}_{\rho,12}^{\mathrm{T}}=\frac{1}{2}\kappa_{\mathrm{FM}}\boldsymbol{J}_{\mathrm{FM}}\boldsymbol{J}_{\mathrm{OA}}^{\dagger}\boldsymbol{\Lambda}_{\mathrm{OA}}$, $\boldsymbol{\Gamma}_{\rho,22}=\kappa_{\mathrm{FM}}\boldsymbol{J}_{\mathrm{FM}}(\boldsymbol{I}_n-\boldsymbol{J}_{\mathrm{OA}}^{\dagger}\boldsymbol{J}_{\mathrm{OA}})\boldsymbol{J}_{\mathrm{FM}}^{\dagger}\boldsymbol{\Lambda}_{\mathrm{FM}}$, $\boldsymbol{\Gamma}_{\rho,31}=\boldsymbol{\Gamma}_{\rho,13}^{\mathrm{T}}=\frac{1}{2}\kappa_{\mathrm{FR}}\boldsymbol{J}_{\mathrm{FR}}\boldsymbol{J}_{\mathrm{OA}}^{\dagger}\boldsymbol{\Lambda}_{\mathrm{OA}}$, $\boldsymbol{\Gamma}_{\rho,23}=\kappa_{\mathrm{FR}}\boldsymbol{J}_{\mathrm{FR}}(\boldsymbol{I}_n-\boldsymbol{J}_{\mathrm{OA}}^{\dagger}\boldsymbol{J}_{\mathrm{OA}})(\boldsymbol{I}_n-\boldsymbol{J}_{\mathrm{FM}}^{\dagger}\boldsymbol{J}_{\mathrm{FM}})\boldsymbol{J}_{\mathrm{FR}}^{\dagger}\boldsymbol{\Lambda}_{\mathrm{FR}}$, $\boldsymbol{\Gamma}_{\rho,32}=$

$\kappa_{\rm FM}\boldsymbol{J}_{\rm FR}(\boldsymbol{I}_n-\boldsymbol{J}_{\rm OA}^{\dagger}\boldsymbol{J}_{\rm OA})\boldsymbol{J}_{\rm FM}^{\dagger}\boldsymbol{\Lambda}_{\rm FM}$, $\boldsymbol{\Gamma}_{\rho,33}=\kappa_{\rm FR}\boldsymbol{J}_{\rm FR}(\boldsymbol{I}_n-\boldsymbol{J}_{\rm OA}^{\dagger}\boldsymbol{J}_{\rm OA})(\boldsymbol{I}_n-\boldsymbol{J}_{\rm FM}^{\dagger}\boldsymbol{J}_{\rm FM})\boldsymbol{J}_{\rm FR}^{\dagger}\boldsymbol{\Lambda}_{\rm FR}$, $\tilde{\boldsymbol{\zeta}}=[\tilde{\boldsymbol{\rho}}_{\rm OA}^{\rm T},\tilde{\boldsymbol{\rho}}_{\rm FM}^{\rm T},\tilde{\boldsymbol{\rho}}_{\rm FR}^{\rm T}]^{\rm T}$. Eq. (S19) satisfies the following inequality:

$$\begin{aligned}\dot{V}_{\rho}\leq&-\underline{\Gamma}_{\rho,11}\left\|\tilde{\boldsymbol{\rho}}_{\rm OA}\right\|^2-\underline{\Gamma}_{\rho,22}\left\|\tilde{\boldsymbol{\rho}}_{\rm FM}\right\|^2-\underline{\Gamma}_{\rho,33}\left\|\tilde{\boldsymbol{\rho}}_{\rm FR}\right\|^2+2\overline{\Gamma}_{\rho,21}\left\|\tilde{\boldsymbol{\rho}}_{\rm OA}\right\|\left\|\tilde{\boldsymbol{\rho}}_{\rm FM}\right\|\\&+2\overline{\Gamma}_{\rho,31}\left\|\tilde{\boldsymbol{\rho}}_{\rm OA}\right\|\left\|\tilde{\boldsymbol{\rho}}_{\rm FR}\right\|+2\overline{\Gamma}_{\rho,32}\left\|\tilde{\boldsymbol{\rho}}_{\rm FM}\right\|\left\|\tilde{\boldsymbol{\rho}}_{\rm FR}\right\|=-\hat{\boldsymbol{\zeta}}^{\rm T}\hat{\boldsymbol{\Gamma}}_{\rho}\hat{\boldsymbol{\zeta}},\end{aligned}\tag{S20}$$

where

$$\hat{\boldsymbol{\Gamma}}_{\rho}=\begin{bmatrix}\kappa_{\rm OA}\underline{\Lambda}_{\rm OA}&-\dfrac{\kappa_{\rm FR}\overline{\Lambda}_{\rm OA}}{2}&-\dfrac{\kappa_{\rm FR}\overline{\Lambda}_{\rm OA}}{2}\\-\dfrac{\kappa_{\rm FM}\overline{\Lambda}_{\rm OA}}{2}&\kappa_{\rm FM}\underline{\Lambda}_{\rm FM}&-\kappa_{\rm FM}\overline{\Lambda}_{\rm FM,FR}\\-\dfrac{\kappa_{\rm FR}\overline{\Lambda}_{\rm OA}}{2}&-\kappa_{\rm FM}\overline{\Lambda}_{\rm FM,FR}&\kappa_{\rm FR}\underline{\Lambda}_{\rm FR}\end{bmatrix}.$$

$\underline{\Gamma}_{\rho,11}=\kappa_{\rm OA}\underline{\Lambda}_{\rm OA}$, $\underline{\Gamma}_{\rho,22}=\kappa_{\rm FM}\underline{\Lambda}_{\rm FM}$, and $\underline{\Gamma}_{\rho,33}=\kappa_{\rm FR}\underline{\Lambda}_{\rm FR}$ are the lower bounds on the induced norms of $\boldsymbol{\Gamma}_{\rho,11}$, $\boldsymbol{\Gamma}_{\rho,22}$, and $\boldsymbol{\Gamma}_{\rho,33}$, respectively; $\underline{\Lambda}_{\rm OA}=\min_i\{\Lambda_{{\rm OA},i}\}$, $\underline{\Lambda}_{\rm FM}=\min_i\{\Lambda_{{\rm FM},i}\}$ $\underline{\Lambda}_{\rm FR}=\min_i\{\Lambda_{{\rm FR},i}\}$; $\overline{\Gamma}_{\rho,21}=\kappa_{\rm FM}\overline{\Lambda}_{\rm OA}$, $\overline{\Gamma}_{\rho,31}=\kappa_{\rm FR}\overline{\Lambda}_{\rm OA}$, and $\overline{\Gamma}_{\rho,32}=2\kappa_{\rm FM}\overline{\Lambda}_{\rm FM,FR}$ are the upper bounds on the induced norms of $\boldsymbol{\Gamma}_{\rho,21}$, $\boldsymbol{\Gamma}_{\rho,31}$, and $\boldsymbol{\Gamma}_{\rho,32}$, respectively; $\overline{\Lambda}_{\rm OA}=\max_i\{\Lambda_{{\rm OA},i}\}$, $\overline{\Lambda}_{\rm FM,FR}=\max_i\{\Lambda_{{\rm FM},i},\Lambda_{{\rm FR},i}\}$, $\hat{\boldsymbol{\zeta}}=\left[\left\|\tilde{\boldsymbol{\rho}}_{\rm OA}\right\|,\left\|\tilde{\boldsymbol{\rho}}_{\rm FM}\right\|,\left\|\tilde{\boldsymbol{\rho}}_{\rm FR}\right\|\right]^{\rm T}$, $\left\|\boldsymbol{I}_n-\boldsymbol{J}_{\rm OA}^{\dagger}\boldsymbol{J}_{\rm OA}\right\|\leq 1$, $\left\|\boldsymbol{I}_n-\boldsymbol{J}_{\rm FM}^{\dagger}\boldsymbol{J}_{\rm FM}\right\|\leq 1$. When

$$\kappa_{\rm OA}>\max\left\{\frac{\kappa_{\rm FM}\overline{\Lambda}_{\rm OA}^2}{4\underline{\Lambda}_{\rm OA}\underline{\Lambda}_{\rm FM}},\frac{\kappa_{\rm FR}\overline{\Lambda}_{\rm OA}\left(2\kappa_{\rm FM}\overline{\Lambda}_{\rm FM,FR}+\kappa_{\rm FR}\underline{\Lambda}_{\rm FM}+\kappa_{\rm FM}\underline{\Lambda}_{\rm FR}\right)}{4\left(\kappa_{\rm FR}\underline{\Lambda}_{\rm FM}\underline{\Lambda}_{\rm FR}-\kappa_{\rm FM}\overline{\Lambda}_{\rm FM,FR}^2\right)}\right\},$$

we may infer that $\hat{\boldsymbol{\Gamma}}_{\rho}$ is a positive-definite symmetric matrix,

$$\dot{V}_{\rho}\leq-\underline{\lambda}_{\hat{\Gamma}}\hat{\boldsymbol{\zeta}}^{\rm T}\hat{\boldsymbol{\zeta}}\leq 0,\tag{S21}$$

where $\underline{\lambda}_{\hat{\Gamma}}$ is the minimum eigenvalue of $\hat{\boldsymbol{\Gamma}}_{\rho}$. Other behavior priority cases are proved by similar ways. Since Assumption 3 is held, control inputs are not always saturated, and mission errors are converged eventually.

We consider a special case in which the FM and FR behaviors are conflicting, and the OA behavior is conflict-free. We also assume that the priority is FM>FR. Then, the differential of Eq. (S18) is expressed as

$$\dot{V}_{\rho}=-\kappa_{\rm OA}\tilde{\boldsymbol{\rho}}_{\rm OA}^{\rm T}\boldsymbol{J}_{\rm OA}\boldsymbol{J}_{\rm OA}^{\dagger}\boldsymbol{\Lambda}_{\rm OA}\tilde{\boldsymbol{\rho}}_{\rm OA}-\tilde{\boldsymbol{\zeta}}^{\rm T}\boldsymbol{\Gamma}_{\rho}\tilde{\boldsymbol{\zeta}},\tag{S22}$$

where $\tilde{\boldsymbol{\zeta}}=[\tilde{\boldsymbol{\rho}}_{\rm FM}^{\rm T},\tilde{\boldsymbol{\rho}}_{\rm FR}^{\rm T}]^{\rm T}$, $\boldsymbol{\Gamma}_{\rho}=\begin{bmatrix}\boldsymbol{\Gamma}_{\rho,11}&\boldsymbol{\Gamma}_{\rho,12}\\\boldsymbol{\Gamma}_{\rho,21}&\boldsymbol{\Gamma}_{\rho,22}\end{bmatrix}$, $\boldsymbol{\Gamma}_{\rho,11}=\kappa_{\rm FM}\boldsymbol{\Lambda}_{\rm FM}$, $\boldsymbol{\Gamma}_{\rho,21}=\boldsymbol{\Gamma}_{\rho,12}^{\rm T}=\frac{1}{2}\kappa_{\rm FR}\boldsymbol{J}_{\rm FR}\boldsymbol{J}_{\rm FM}^{\dagger}\boldsymbol{\Lambda}_{\rm FM}$, and $\boldsymbol{\Gamma}_{\rho,22}=\kappa_{\rm FR}\boldsymbol{J}_{\rm FM}\left(\boldsymbol{I}-\boldsymbol{J}_{\rm FM}\boldsymbol{J}_{\rm FM}^{\dagger}\right)\boldsymbol{J}_{\rm FR}^{\dagger}\boldsymbol{\Lambda}_{\rm FR}$. Eq. (S22) satisfies the following inequality:

$$\dot{V}_{\rho}\leq-\underline{\Lambda}_{\rm OA}\left\|\tilde{\boldsymbol{\rho}}_{\rm OA}\right\|^2-\underline{\Gamma}_{\rho,11}\left\|\tilde{\boldsymbol{\rho}}_{\rm FM}\right\|^2-\underline{\Gamma}_{\rho,22}\left\|\tilde{\boldsymbol{\rho}}_{\rm FR}\right\|^2+2\overline{\Gamma}_{\rho,21}\left\|\tilde{\boldsymbol{\rho}}_{\rm FM}\right\|\left\|\tilde{\boldsymbol{\rho}}_{\rm FR}\right\|=-\hat{\boldsymbol{\zeta}}^{\rm T}\hat{\boldsymbol{\Gamma}}_{\rho}\hat{\boldsymbol{\zeta}},\tag{S23}$$

where $\underline{\Gamma}_{\rho,11}$ and $\underline{\Gamma}_{\rho,22}$ are the lower bounds on the induced norms of $\boldsymbol{\Gamma}_{\rho,11}$ and $\boldsymbol{\Gamma}_{\rho,22}$ respectively, $\underline{\Gamma}_{\rm FM}=\min_i\{\Gamma_{{\rm FM},i}\}$, $\underline{\Gamma}_{\rm FR}=\min_i\{\Gamma_{{\rm FR},i}\}$, $\overline{\Gamma}_{\rho,21}$ is the upper bound on the induced norm of $\boldsymbol{\Gamma}_{\rho,12}$, $\overline{\Gamma}_{\rm FM}=\max_i\{\Gamma_{{\rm FM},i}\}$, $\hat{\boldsymbol{\zeta}}=\left[\left\|\tilde{\boldsymbol{\rho}}_{\rm FM}\right\|^{\rm T},\left\|\tilde{\boldsymbol{\rho}}_{\rm FR}\right\|^{\rm T}\right]^{\rm T}$, $\hat{\boldsymbol{\Gamma}}_{\rho}=\begin{bmatrix}\hat{\Gamma}_{\rho,11}&\hat{\Gamma}_{\rho,12}\\\hat{\Gamma}_{\rho,21}&\hat{\Gamma}_{\rho,22}\end{bmatrix}$, $\hat{\Gamma}_{\rho,11}=\kappa_{\rm FM}\underline{\Lambda}_{\rm FM}$, $\hat{\Gamma}_{\rho,12}=\hat{\Gamma}_{\rho,21}=-\kappa_{\rm FR}\overline{\Lambda}_{\rm FR}/2$, and $\hat{\Gamma}_{\rho,22}=\kappa_{\rm FR}\underline{\Lambda}_{\rm FR}$. When $\kappa_{\rm FM}>\dfrac{\kappa_{\rm FR}\overline{\Lambda}_{\rm FM}^2}{4\overline{\Lambda}_{\rm FM}\overline{\Lambda}_{\rm FR}}$, $\hat{\boldsymbol{\Gamma}}_{\rho}$ is a positive-definite symmetric matrix,

$$\dot{V}_{\rho}\leq-\underline{\Lambda}_{\rm OA}\left\|\tilde{\boldsymbol{\rho}}_{\rm OA}\right\|^2-\underline{\lambda}_{\hat{\Gamma}}\hat{\boldsymbol{\zeta}}^{\rm T}\hat{\boldsymbol{\zeta}}\leq 0,\tag{S24}$$

where $\underline{\lambda}_{\hat{\Gamma}}$ is the minimum eigenvalue of $\hat{\boldsymbol{\Gamma}}_{\rho}$. The proof is completed.

## 3 Proof of boundedness

Since inequality (S17) only gives the total bound of all error signals, the bounds of each error signal should be further analyzed. Let us define the Lyapunov function candidates as $V_\tau = \frac{1}{2}\sum_{i=1}^{N}\boldsymbol{\tau}_i^{\mathrm{T}}\boldsymbol{\tau}_i$, $V_{\mathrm{e}} = \frac{1}{2}\sum_{i=1}^{N}\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{\Upsilon}_{\mathrm{e},i}\boldsymbol{e}_i$, $V_{\mathrm{f}} = \frac{1}{2}\sum_{i=1}^{N}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}^{\mathrm{T}}\boldsymbol{\Upsilon}_{\mathrm{f},i}^{-1}\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}\right\}$, $V_{\mathrm{c}} = \frac{1}{2}\sum_{i=1}^{N}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}}\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}\right\}$, and $V_{\mathrm{a}} = \frac{1}{2}\sum_{i=1}^{N}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}$. Following the above proof, it is easy to yield the results as

$$\dot{V}_\tau \leq -\sum_{i=1}^{N}\left(\xi_{\tau,i} - \frac{1}{2}\right)\|\boldsymbol{\tau}_i\| + \frac{N}{2}\epsilon_{\Delta}. \tag{S25}$$

Let $\vartheta_\tau = 2\xi_{\tau,i} - 1$ and $\vartheta_{\Delta} = \frac{N}{2}\epsilon_{\mathrm{f}}$; we then have

$$\dot{V}_\tau \leq -\vartheta_\tau V_\tau + \vartheta_{\Delta}. \tag{S26}$$

The inequality is obtained by using Lemma 1 as

$$V_\tau \leq \mathrm{e}^{-\vartheta_\tau t}V_\tau(0) + \frac{\vartheta_{\Delta}}{\vartheta_\tau}(1 - \mathrm{e}^{-\vartheta_\tau t}), \tag{S27}$$

where $\boldsymbol{\tau}_i$ is bounded with $\|\boldsymbol{\tau}_i\| \leq \epsilon_\tau = \left[2\mathrm{e}^{-\vartheta_\tau t}V_\tau(0) + \frac{2\vartheta_{\Delta}}{\vartheta_\tau}(1 - \mathrm{e}^{-\vartheta_\tau t})\right]^{1/2}$.

$$\dot{V}_{\mathrm{e}} + \dot{V}_{\mathrm{f}} \leq -\frac{1}{2}\sum_{i=1}^{N}\frac{\underline{\lambda}_\xi}{\overline{\lambda}_{\mathrm{e}}}\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{\Upsilon}_{\mathrm{e},i}\boldsymbol{e}_i - \frac{1}{2}\sum_{i=1}^{N}\frac{\nu_{\mathrm{f},i}}{\overline{\lambda}_{\mathrm{f}}}\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}^{\mathrm{T}}\boldsymbol{\Upsilon}_{\mathrm{f},i}^{-1}\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}\right\} + \Xi_\xi, \tag{S28}$$

where $\Xi_\xi = \sum_{i=1}^{N}\|\dot{\boldsymbol{v}}_{i,\mathrm{r}}\|^2 + N\varepsilon_{\mathrm{f}}^2 + \epsilon_\tau + \frac{1}{4\beta_{\mathrm{V}}^2}\sum_{i=1}^{N}\mathrm{tr}\left\{\hat{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}$. Let $\vartheta_{\mathrm{e}} = \frac{\underline{\lambda}_\xi}{\overline{\lambda}_{\mathrm{e}}}$, $\vartheta_{\mathrm{f}} = \frac{\nu_{\mathrm{f},i}}{\overline{\lambda}_{\mathrm{f},i}}$, and $|\Xi_\xi| \leq \vartheta_\xi$; we then have

$$\dot{V}_{\mathrm{e}} \leq -\vartheta_{\mathrm{e}}V_{\mathrm{e}} + \vartheta_\xi, \tag{S29}$$

$$\dot{V}_{\mathrm{f}} \leq -\vartheta_{\mathrm{f}}V_{\mathrm{f}} + \vartheta_\xi. \tag{S30}$$

The inequalities are obtained by using Lemma 1 as

$$V_{\mathrm{e}} \leq \mathrm{e}^{-\vartheta_{\mathrm{e}}t}V_{\mathrm{e}}(0) + \frac{\vartheta_\xi}{\vartheta_{\mathrm{e}}}(1 - \mathrm{e}^{-\vartheta_{\mathrm{e}}t}), \tag{S31}$$

$$V_{\mathrm{f}} \leq \mathrm{e}^{-\vartheta_{\mathrm{f},1}t}V_{\mathrm{f}}(0) + \frac{\vartheta_\xi}{\vartheta_{\mathrm{f}}}(1 - \mathrm{e}^{-\vartheta_{\mathrm{f}}t}), \tag{S32}$$

where $\boldsymbol{e}_i$ is bounded with $\|\boldsymbol{e}_i\| \leq \epsilon_{\mathrm{e}} = \left[2\mathrm{e}^{-\vartheta_{\mathrm{e}}t}V_{\mathrm{e}}(0) + \frac{2\vartheta_\xi}{\vartheta_{\mathrm{e}}}(1 - \mathrm{e}^{-\vartheta_{\mathrm{e}}t})\right]^{1/2}$, $\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}$ is bounded with $\|\tilde{\boldsymbol{\omega}}_{\mathrm{f},i}\| \leq \epsilon_{\mathrm{f}} = \left\{\left[2\mathrm{e}^{-\vartheta_{\mathrm{f}}t}V_{\mathrm{f}}(0) + \frac{2\vartheta_\xi}{\vartheta_{\mathrm{f}}}(1 - \mathrm{e}^{-\vartheta_{\mathrm{f}}t})\right]/\underline{\lambda}_{\mathrm{f}}\right\}^{1/2}$, $\underline{\lambda}_{\mathrm{f}}$ is the minimal eigenvalue of $\boldsymbol{\Upsilon}_{\mathrm{f},i}^{-1}$.

$$\dot{V}_{\mathrm{c}} + \dot{V}_{\mathrm{a}} \leq -\frac{1}{2}\sum_{i=1}^{N}\underline{\lambda}_\Psi\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}^{\mathrm{T}}\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}\right\} - \frac{1}{2}\sum_{i=1}^{N}\underline{\lambda}_\Psi\mathrm{tr}\left\{\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}\right\} + \Xi_\Psi, \tag{S33}$$

where $\Xi_\Psi = \frac{1}{4\beta_{\mathrm{V}}^2}\sum_{i=1}^{N}\mathrm{tr}\left\{\hat{\boldsymbol{\omega}}_{\mathrm{a},i}^{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{V},i}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\boldsymbol{\Psi}_{\mathrm{V},i}^{\mathrm{T}}(\boldsymbol{\chi}_i,\boldsymbol{\tau}_i)\hat{\boldsymbol{\omega}}_{\mathrm{a},i}\right\}$ with $|\Xi_\Psi| \leq \vartheta_\Psi$. Let $\vartheta_{\mathrm{c},1} = \underline{\lambda}_\Psi$ and $\vartheta_{\mathrm{a},1} = \underline{\lambda}_\Psi$; we then have

$$\dot{V}_{\mathrm{c}} \leq -\vartheta_{\mathrm{c}}V_{\mathrm{c}} + \vartheta_\Psi, \tag{S34}$$

$$\dot{V}_{\mathrm{a}} \leq -\vartheta_{\mathrm{a}}V_{\mathrm{a}} + \vartheta_\Psi. \tag{S35}$$

The inequalities are obtained by using Lemma 1 as

$$V_{\mathrm{c}} \leq \mathrm{e}^{-\vartheta_{\mathrm{c}} t} V_{\mathrm{c}}(0) + \frac{\vartheta_{\Psi}}{\vartheta_{\mathrm{c}}}(1 - \mathrm{e}^{-\vartheta_{\mathrm{c}} t}), \tag{S36}$$

$$V_{\mathrm{a}} \leq \mathrm{e}^{-\vartheta_{\mathrm{a},1} t} V_{\mathrm{a}}(0) + \frac{\vartheta_{\Psi}}{\vartheta_{\mathrm{a}}}(1 - \mathrm{e}^{-\vartheta_{\mathrm{a}} t}), \tag{S37}$$

where $\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}$ is bounded with $\|\tilde{\boldsymbol{\omega}}_{\mathrm{c},i}\| \leq \epsilon_{\mathrm{c}} = \left[2\mathrm{e}^{-\vartheta_{\mathrm{c}} t} V_{\mathrm{c}}(0) + \frac{2\vartheta_{\Psi}}{\vartheta_{\mathrm{c}}}(1 - \mathrm{e}^{-\vartheta_{\mathrm{c}} t})\right]^{1/2}$ and $\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}$ is bounded with $\|\tilde{\boldsymbol{\omega}}_{\mathrm{a},i}\| \leq \epsilon_{\mathrm{a}} = \left[2\mathrm{e}^{-\vartheta_{\mathrm{a}} t} V_q(0) + \frac{2\vartheta_{\Psi}}{\vartheta_{\mathrm{a}}}(1 - \mathrm{e}^{-\vartheta_{\mathrm{a}} t})\right]^{1/2}$.

**Algorithm S1** MARLMS

**Input:** total number of training episodes $T_{\mathrm{e}}$, total number of time steps $T_{\mathrm{s}}$

Initialize $Q(\boldsymbol{s}_t, \boldsymbol{b}_t; \omega_{\mathrm{Q}}, \omega_{\mathrm{V}}, \omega_{\mathrm{B}}) = V(\boldsymbol{s}_t; \omega_{\mathrm{Q}}, \omega_{\mathrm{V}}) + B(\boldsymbol{s}_t, \boldsymbol{b}_t; \omega_{\mathrm{Q}}, \omega_{\mathrm{B}})$ with random weights

Initialize experience replay buffer $\mathfrak{D}$

Initialize $\bar{T}(\phi(\boldsymbol{s}_t))$-greedy exploration and leniency $\mathfrak{L}(\boldsymbol{s}_t, \boldsymbol{b}_t)$

**For** episode$=1:T_{\mathrm{e}}$ **do**

Reset the joint state $\boldsymbol{s}_t$ to initial value $\boldsymbol{s}_0$

  **For** $t = 1 : T_{\mathrm{s}}$ **do**

  $Q(\boldsymbol{s}_{t+1}, \boldsymbol{b}_{t+1}; \omega_{\mathrm{Q}}^{-}, \omega_{\mathrm{V}}^{-}, \omega_{\mathrm{B}}^{-}) = \frac{1}{\lambda} \sum_{\iota=1}^{\lambda} Q(\boldsymbol{s}_{t+1}, \boldsymbol{b}_{t+1}; \omega_{Q_{t-\iota}}^{-}, \omega_{V_{t-\iota}}^{-}, \omega_{B_{t-\iota}}^{-})$

  $y_{\boldsymbol{s}_t, \boldsymbol{b}_t} = \mathbb{E}_D[r + \gamma \max_{\boldsymbol{b}_{t+1}} Q(\boldsymbol{s}_{t+1}, \boldsymbol{b}_{t+1}; \omega_{\mathrm{Q}}^{-}, \omega_{\mathrm{V}}^{-}, \omega_{\mathrm{B}}^{-}) | \boldsymbol{s}_t, \boldsymbol{b}_t]$

  $$\begin{cases} \omega_{Q_t}, \omega_{V_t}, \omega_{B_t} \approx \arg\min_{\omega_{\mathrm{Q}}, \omega_{\mathrm{V}}, \omega_{\mathrm{B}}} \mathbb{E}_D[(y_{\boldsymbol{s}_t, \boldsymbol{b}_t} - Q(\boldsymbol{s}_t, \boldsymbol{b}_t; \omega_{\mathrm{Q}}, \omega_{\mathrm{V}}, \omega_{\mathrm{B}}))]^2, & \delta_t > 0 \text{ or } \vartheta > \mathfrak{L}_t \\ \omega_{Q_t}, \omega_{V_t}, \omega_{B_t} = \omega_{Q_t}, \omega_{V_t}, \omega_{B_t}, & \delta_t \leq 0 \text{ and } \vartheta \leq \mathfrak{L}_t \end{cases}$$

  Update $\mathfrak{D}$, $\bar{T}(\phi(\boldsymbol{s}_t))$, and $\mathfrak{L}_t$

  **End for**

**End for**

**Output:** $Q(\boldsymbol{s}_t, \boldsymbol{b}_t; \omega_{\mathrm{Q}}, \omega_{\mathrm{V}}, \omega_{\mathrm{B}})$