Neural Network + RISE Control Design

December 17, 2021

1 Problem Statement

System:

$$\begin{cases} \dot{x}_1 & = x_2 \\ M_w(x_1)\dot{x}_2 & = f_{1w}(x) + x_3 + d_{1w}(x,t) \\ \dot{x}_3 & = x_4 \\ J_w\dot{x}_4 & = f_{2w}(x) + u + d_{2w}(t) \end{cases}, \tag{1}$$

where $x = \begin{bmatrix} x_1^\top & x_2^\top & x_3^\top & x_4^\top \end{bmatrix}^\top \in \mathbb{R}^8$ is the system state, functions $f_{iw}(x) \in \mathbb{R}^2$ and $M_w(x_1) \in \mathbb{R}^{2\times 2} > 0$ represent unknown dynamics with continuously differentiable nominal models $\bar{f}_{iw}(x)$ and $\bar{M}_w(x_1)$, respectively, $J_w \in \mathbb{R}^{2\times 2}$ is an unknown constant parameter, $d_{iw} \in \mathbb{R}^2$ are continuous transformed time-varying disturbance such that $\|d_{1w}(x,t)\|_1 \le \rho_0$ and $\left(d_{2w}(t), \dot{d}_{2w}(t)\right) \in \mathcal{L}_\infty$ with some known constant ρ_0.

The control objective is to design a controller $u(t)$ to ensure a practically stable tracking of $x_1(t)$ to some known trajectory $x_{1d}(t)$ satisfying $(x_{1d}, \dot{x}_{1d}, \ddot{x}_{1d}) \in \mathcal{L}_\infty$.

2 Control Formulation

We apply the Backstepping Method as follows.

Step 1:

Define the tracking error

$$e_1(t) = x_1(t) - x_{1d}(t), \tag{2}$$

and set

$$s = \dot{e}_1 + \sigma_1 e_1, \tag{3}$$

we have

$$M_w(x_1)\dot{s} = f_{1w}(x) + x_3 + d_{1w}(x,t) + M_w(x_1)\left(\sigma_1\dot{e}_1 - \ddot{x}_{1d}\right). \tag{4}$$

Define

$$z_1(t) = x_3(t) - p(t), \tag{5}$$

where $p(t) \in \mathbb{R}^2$ is the output of the following 3-rd order filter

$$\begin{cases} \dot{\xi} & = A\xi + BC\alpha \\ p & = C\xi \end{cases}, \tag{6}$$

where $\xi(t) = \begin{bmatrix} \xi_1^\top(t) & \xi_2^\top(t) & \xi_3^\top(t) \end{bmatrix}^\top \in \mathbb{R}^6$ is the filter state, $\alpha(t) = \begin{bmatrix} \alpha_1^\top(t) & \dot{\alpha}_1^\top(t) & \ddot{\alpha}_1^\top(t) \end{bmatrix}^\top \in \mathbb{R}^6$ with $\alpha_1(t) \in \mathbb{R}^2$ being the virtual control to be designed, the matrix

$$A = \begin{bmatrix} \mathbf{0}_{2\times 2} & I_{2\times 2} & \mathbf{0}_{2\times 2} \\ \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times 2} & I_{2\times 2} \\ \boldsymbol{a_1} & \boldsymbol{a_2} & \boldsymbol{a_3} \end{bmatrix},\ B = \begin{bmatrix} \mathbf{0}_{2\times 2} \\ \mathbf{0}_{2\times 2} \\ -\boldsymbol{a_1} \end{bmatrix},\ C = \begin{bmatrix} I_{2\times 2} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times 2} \end{bmatrix} \tag{7}$$

with $\boldsymbol{a_i} = a_i I_{2\times 2}$ and $a_i < 0$ such that A is Hurwitz.

Define

$$y(t) = \xi(t) - \alpha(t), \tag{8}$$

we have

$$\dot{y} = Ay + (A + BC)\,\alpha - \dot{\alpha}. \tag{9}$$

Using (5)-(8), we know that

$$x_3 = z_1 + Cy + \alpha_1. \tag{10}$$

Define $\underline{\bullet} = \bullet - \bar{\bullet}$ as the uncertain dynamics and substitute (10) into (4), we obtain

$$\begin{aligned} M_w(x_1)\dot{s} = {} & \bar{F}_1(x,s,t) + z_1 + Cy + \alpha_1 + D_1(x,s,t) + d_{1w}(x,t) \\ & - \frac{1}{2}\dot{\bar{M}}_w(x_1)s - \frac{1}{2}\dot{\underline{M}}_w(x_1)s, \end{aligned} \tag{11}$$

where

$$\begin{cases} \bar{F}_1 x, s, t & = \bar{f}_{1w}(x) + \bar{M}_w(x_1)\left(\sigma_1 \dot{e}_1 - \ddot{x}_{1d}\right) + \frac{1}{2}\dot{\bar{M}}_w(x_1)s \\ D_1(x,s,t) & = \underline{f}_{1w}(x) + \underline{M}_w(x_1)\left(\sigma_1 \dot{e}_1 - \ddot{x}_{1d}\right) + \frac{1}{2}\dot{\underline{M}}_w(x_1)s, \end{cases} \tag{12}$$

and $\dot{M}_w(x_1) = \dot{\bar{M}}_w(x_1) + \dot{\underline{M}}_w(x_1) = \frac{\partial M_w(x_1)}{\partial x_1} x_2$. Note that $D_1(x,s,t)$ is the unknown dynamics without explicit t from any unknown time-varying functions, we are encouraged to apply RBFNN to approximate it by

$$D_1(x,s,t) = W^\top \Lambda\left(\bar{Z}\right) + \delta\left(\bar{Z}\right), \tag{13}$$

with $W \in \mathbb{R}^{l\times 2}$ being the ideal output layer weight and $\Lambda\left(\bar{Z}\right) = \begin{bmatrix} \lambda_1\left(\bar{Z}\right) & \lambda_2\left(\bar{Z}\right) & \cdots & \lambda_l\left(\bar{Z}\right) \end{bmatrix}^\top$, where $l > 0$ is the neuro number of the hidden layer, and the NN input vector $\bar{Z} = \begin{bmatrix} Z^\top & 1 \end{bmatrix}^\top$ with $Z = \begin{bmatrix} x^\top & \dot{e}_1^\top & \ddot{x}_{1d}^\top & s^\top \end{bmatrix}^\top \in \mathbb{R}^8$, each $\lambda_i\left(\bar{Z}\right)$ is the Gaussian function

$$\lambda_i\left(\bar{Z}\right) = e^{-\frac{\left(\bar{Z}-c\right)^\top\left(\bar{Z}-c\right)}{2b^2}} \tag{14}$$

with $c \in \mathbb{R}^9$ and $b \in \mathbb{R}^9$ being the center and the shape factor, respectively, $\delta\left(Z\right) \in \mathbb{R}^2$ is the NN approximation error such that

$$\left\|\delta\left(Z\right)\right\|_2 \le \bar{\delta} \tag{15}$$

with $\bar{\delta} > 0$ being some unknown constant.

Design the NN virtual control

$$\alpha_1 = -\bar{F}_1(x,s,t) - K_1 s - \hat{W}^\top(t)\Lambda\left(\bar{Z}\right) - \rho \mathrm{Tanh}\left(\frac{s}{\epsilon}\right), \tag{16}$$

where $K_1 \in \mathbb{R}^{2\times 2} > 0$, $\epsilon > 0$ and $\rho \ge \rho_0 + \bar{\delta}$ are control gains, $\hat{W}(t) \in \mathbb{R}^{l\times 2}$ is the online estimate of the unknown NN weight W, and $\mathrm{Tanh}\left(s/\epsilon\right) = \begin{bmatrix} \tanh\left(s_1/\epsilon\right) & \tanh\left(s_2/\epsilon\right) \end{bmatrix}^\top$. Substituting (16) into (11) gives

$$\begin{aligned} M_w(x_1)\dot{s} = {} & -K_1 s - \tilde{W}^\top(t)\Lambda\left(\bar{Z}\right) + \left(\delta\left(Z\right) + d_{1w}(x,t) - \rho\mathrm{Tanh}\left(\frac{s}{\epsilon}\right)\right) \\ & + z_1 + Cy - \frac{1}{2}\dot{\bar{M}}_w(x_1)s - \frac{1}{2}\dot{\underline{M}}_w(x_1)s \end{aligned} \tag{17}$$

where $\tilde{W}(t) = \hat{W}(t) - W$ is the NN weight estimate error.

Define a positive definite function

$$V_1 = \frac{1}{2}s^\top M_w(x_1)s + \frac{1}{2\gamma}\mathrm{tr}\left\{\tilde{W}^\top \Gamma^{-1}\tilde{W}\right\} + \frac{1}{2}z_1^\top z_1 + y^\top P y, \tag{18}$$

where $\gamma > 0$ and $\Gamma = \Gamma^\top \in \mathbb{R}^{l\times l} > 0$ are control gains, and $P = P^\top \in \mathbb{R}^{6\times 6} > 0$ is some constant matrix. Taking time derivative of (18) yields

$$\begin{aligned} \dot{V}_1 = {} & -s^\top K_1 s + s^\top\left(\delta\left(Z\right) + d_{1w}(x,t) - \rho\mathrm{Tanh}\left(\frac{s}{\epsilon}\right)\right) + s^\top\left(z_1 + Cy\right) \\ & + \frac{1}{\gamma}\mathrm{tr}\left\{\tilde{W}^\top\left(\Gamma^{-1}\dot{\hat{W}} - \gamma\Lambda\left(\bar{Z}\right)s^\top\right)\right\} + z_1^\top \dot{z}_1 + \dot{y}^\top P y + y^\top P \dot{y}. \end{aligned} \tag{19}$$

It is known that

$$s^\top \left(\delta\left(Z\right) + d_{1w}(x,t) - \rho \mathrm{Tanh}\left(\frac{s}{\epsilon}\right)\right) \le \rho \sum_{i=1}^{2} \left(|s_i| - s_i \tanh\left(\frac{s_i}{\epsilon}\right)\right) \le 2\rho c_0 \epsilon \tag{20}$$

with $c_0 \approx 0.2875$, and by Young's Inequality, we also have

$$s^\top \left(z_1 + Cy\right) \le \left(\frac{1}{4\epsilon_1^2} + \frac{1}{4\epsilon_2^2}\right) s^\top s + \epsilon_1^2 z_1^\top z_1 + \epsilon_2^2 y^\top C^\top C y, \tag{21}$$

where ϵ_i are any real numbers. Based on (9), we also have the following relation

$$\begin{aligned} \dot{y}^\top P y + y^\top P \dot{y} = & \left[y^\top A^\top + \alpha^\top \left(A^\top + C^\top B^\top\right) - \dot{\alpha}^\top\right] P y \\ & + y^\top P \left[Ay + \left(A + BC\right)\alpha - \dot{\alpha}\right] \\ = & \, y^\top \left(A^\top P + PA\right) y + \left[\alpha^\top \left(A^\top + C^\top B^\top\right) P y + y^\top P \left(A + BC\right)\alpha\right] \\ & - \left(\dot{\alpha}^\top P y + y^\top P \dot{\alpha}\right), \end{aligned} \tag{22}$$

To find a proper positive definite matrix P, we require that

$$\dot{y}^\top P y + y^\top P \dot{y} + \epsilon_2^2 y^\top C^\top C y \le -y^\top Q y, \tag{23}$$

where $Q = Q^\top \in \mathbb{B}^6 > 0$ is any positive definite matrix. To make (23) more feasible, we instead require that

$$\dot{y}^\top P y + y^\top P \dot{y} + \epsilon_2^2 y^\top C^\top C y \le -y^\top Q y + \kappa_1 \alpha^\top \alpha + \kappa_2 \dot{\alpha}^\top \dot{\alpha} \tag{24}$$

with $\kappa_1, \kappa_2 > 0$. Therefore, we rewrite (22) as

$$\dot{y}^\top P y + y^\top P \dot{y} + y^\top \left(\epsilon_2^2 C^\top C + Q\right) y - \kappa_1 \alpha^\top \alpha - \kappa_2 \dot{\alpha}^\top \dot{\alpha} = Y^\top \Phi Y \le 0, \tag{25}$$

where $Y = \begin{bmatrix} y^\top & \alpha^\top & \dot{\alpha}^\top \end{bmatrix}^\top$, and

$$\Phi = \begin{bmatrix} A^\top P + PA + \epsilon_2^2 C^\top C + Q & P\left(A + BC\right) & -P \\ \star & -\kappa_1 I_{6\times 6} & \mathbf{0}_{6\times 6} \\ \star & \star & -\kappa_2 I_{6\times 6} \end{bmatrix} \le 0. \tag{26}$$

Solving the LMI (26) along with $P > 0$, $Q > 0$ and some constant ϵ_2 is capable of providing us with the required matrix P. By (20), (21) and (24), we now can rewrite (19) as

$$\begin{aligned} \dot{V}_1 \le & -s^\top \left(K_1 - \left(\frac{1}{4\epsilon_1^2} + \frac{1}{4\epsilon_2^2}\right) I\right) s + z_1^\top \left(\dot{z}_1 + \epsilon_1^2 z_1\right) - y^\top Q y + 2\rho c_0 \epsilon \\ & + \frac{1}{\gamma} \mathrm{tr}\left\{\tilde{W}^\top \left(\Gamma^{-1} \dot{\hat{W}} - \gamma \Lambda\left(\bar{Z}\right) s^\top\right)\right\} + \kappa_1 \alpha^\top \alpha + \kappa_2 \dot{\alpha}^\top \dot{\alpha}. \end{aligned} \tag{27}$$

.

Step 2:

A RISE-typed controller is designed at this step. First we define two new filtered errors

$$z_2 = \dot{z}_1 + \sigma_2 z_1 \tag{28}$$

and

$$r = \dot{z}_2 + \sigma_3 z_2 + z_1, \tag{29}$$

where $\sigma_i > 0$ are control gains. According to (28)-(29), it is clear that

$$\dot{z}_1 = z_2 - \sigma_2 z_1 \tag{30}$$

and

$$\dot{z}_2 = r - \sigma_3 z_2 - z_1, \tag{31}$$

We use the last equation of (1) and (29) to derive that the open-loop equation of $r(t)$ is

$$J_w r = u + f_{2w}(x) + d_{2w}(t) + J_w\left(\sigma_2 \dot{z}_1 + \sigma_3 z_2 + z_1 - \xi_3\right). \tag{32}$$

Note that the derivation of (32) uses the facts that $\dot{z}_1 = x_4 - \xi_2$ based on (5) and $\ddot{z}_1 = \dot{x}_4 - \xi_3$. Similar to Step 1, we separate the unknown dynamics in (32) by the nominal models and the uncertain ones as follows:

$$\begin{cases} f_{2w}(x) & = \bar{f}_{2w}(x) + \underline{f}_{2w}(x) \\ J_w & = \bar{J}_w + \underline{J}_w \end{cases} \tag{33}$$

Therefore, we rewrite (32) as

$$J_w r = u_r + S\left(x, \xi, t\right) + \underline{f}_{2w}\left(x_{1d}, \dot{x}_{1d}, \xi_1, \xi_2\right) - \underline{J}_w \xi_3 + d_{2w}(t), \tag{34}$$

where $u_r = u - \bar{f}_{2w}(x) - \bar{J}_w\left(\sigma_2 \dot{z}_1 + \sigma_3 z_2 + z_1 - \xi_3\right)$, and the function

$$S\left(x, \xi, t\right) = \underline{f}_{2w}\left(x_1, x_2, x_3, x_4\right) - \underline{f}_{2w}\left(x_{1d}, \dot{x}_{1d}, \xi_1, \xi_2\right) + \underline{J}_w\left(\sigma_2 \dot{z}_1 + \sigma_3 z_2 + z_1\right). \tag{35}$$

We develop the dynamics of $r(t)$ according to (34) by

$$J_w \dot{r} = \dot{u}_r + \tilde{N}\left(x, \xi, t\right) - z_2 + N_d\left(\xi, \alpha_1, t\right), \tag{36}$$

where

$$\tilde{N}\left(x, \xi, t\right) = \dot{S}\left(x, \xi, t\right) + z_2, \tag{37}$$

and

$$N_d\left(\xi, \alpha_1, t\right) = \dot{\underline{f}}_{2w}\left(x_{1d}, \dot{x}_{1d}, \xi_1, \xi_2\right) - \underline{J}_w \dot{\xi}_3 + \dot{d}_{2w}(t), \tag{38}$$

and design the RISE-typed controller

$$\begin{cases} u_r & = -K_s z_2 + K_s z_2(0) - \int_0^t K_s\left(\sigma_3 z_2\left(\tau\right) + z_1\left(\tau\right)\right) + \beta \operatorname{sgn}\left(z_2\left(\tau\right)\right) d\tau \\ u & = u_r + \bar{f}_{2w}(x) + \bar{J}_w\left(\sigma_2 \dot{z}_1 + \sigma_3 z_2 + z_1 - \xi_3\right) \end{cases}. \tag{39}$$

Substituting (39) into (36) hence gives

$$J_w \dot{r} = -K_s r + \tilde{N}\left(x, \xi, t\right) - z_2 + \left(N_d\left(\xi, \alpha_1, t\right) - \beta \operatorname{sgn}\left(z_2\right)\right). \tag{40}$$

Now, we define a new positive definite function

$$V = V_1 + \frac{1}{2} z_2^\top z_2 + \frac{1}{2} r^\top J_w r + H(t). \tag{41}$$

where

$$H(t) = \xi_b - \int_0^t N_d(\tau) - \beta \operatorname{sgn}\left(z_2(\tau)\right) d\tau \geq 0 \tag{42}$$

with $\xi_b = \sum_{i=1}^2 \left(\beta\left|z_{2i}(0)\right| - z_{2i}(0) N_{di}(0)\right)$ and the parameter $\beta \geq \|N_d\|_\infty + \frac{1}{\sigma_3}\left\|\dot{N}_d\right\|_\infty$. Time derivative of (41) along (40) with (2) and (30)-(31) simply yields

$$\begin{aligned} \dot{V} \leq & -s^\top\left(K_1 - \left(\frac{1}{4\epsilon_1^2} + \frac{1}{4\epsilon_2^2}\right) I\right) s - \left(\sigma_2 - \epsilon_1^2\right) z_1^\top z_1 - \sigma_3 z_2^\top z_2 + 2\rho c_0 \epsilon \\ & - y^\top Q y - r^\top K_s r + r^\top \tilde{N} + + \frac{1}{\gamma} \operatorname{tr}\left\{\tilde{W}^\top\left(\Gamma^{-1} \dot{\hat{W}} - \gamma \Lambda\left(\bar{Z}\right) s^\top\right)\right\} \\ & + \kappa_1 \alpha^\top \alpha + \kappa_2 \dot{\alpha}^\top \dot{\alpha}. \end{aligned} \tag{43}$$

Since $\tilde{N}$ in (37) is continuously differentiable, we can show that

$$r^\top \tilde{N} \leq \|r\| \phi\left(\|\iota\|\right)\|\iota\| \leq \phi\left(\|\iota\|\right)\|\iota\|^2, \tag{44}$$

where $\iota = \left[\begin{array}{cccccc} s^\top & z_1^\top & z_2^\top & y^\top & \left\|\tilde{W}\right\| & r^\top \end{array} \right]^\top$, and $\phi(\bullet)$ is some unknown positive definite function. We also design the following adaptive law for the online NN weight adaptation:

$$\dot{\hat{W}} = \gamma\Gamma \left[\Lambda\left(\bar{Z}\right) s^\top - \beta\hat{W}\right], \tag{45}$$

where $\beta > 0$ is the adaptation gain. Now we apply the Young's Inequality

$$-\beta\tilde{W}\hat{W} \leq -\beta\left(1 - \frac{1}{4\varepsilon^2}\right)\tilde{W}^\top\tilde{W} + \beta\varepsilon^2 W^\top W \tag{46}$$

with $\varepsilon > 0$ being some constant, then substitute (44)-(45) into (43) and finally obtain

$$\dot{V} \leq -\left(a_1 - \phi\left(\|\iota\|\right)\right)\|\iota\|^2 + \kappa_1\alpha^\top\alpha + \kappa_2\dot{\alpha}^\top\dot{\alpha} + 2\rho c_0\epsilon + \beta\varepsilon^2 W^\top W, \tag{47}$$

where $a_1 = \min\left\{K_1 - \left(\frac{1}{e\epsilon_1^2} + \frac{1}{4\epsilon_2^2}\right)I, \sigma_2 - \epsilon_1^2, \sigma_3, \lambda_{\min}\left(Q\right), \beta\left(1 - \frac{1}{4\varepsilon_2}\right), \lambda_{\min}\left(K_s\right)\right\}$ with $\lambda_{\min}\left(\bullet\right)$ representing the minimum eigenvalue of the matrix $\bullet$. Using standard boundedness analysis in [1], we know that $\|\iota\|$ is semi-globally uniformly bounded by

$$\|\iota\| \leq \sqrt{\frac{D}{a_1 - \phi\left(\|\iota\|\right)}}, \tag{48}$$

where D is a constant for the boundedness of $\kappa_1\alpha^\top\alpha + \kappa_2\dot{\alpha}^\top\dot{\alpha} + 2\rho c_0\epsilon + \beta\varepsilon^2 W^\top W$ under the semi-global conditions. Note that the control parameter is capable of picking a_1 arbitrarily large. As a result, the boundedness of $\|\iota\|$ can be arbitrarily small. Therefore, the boundedness of $\|s\|$ can be established by

$$\|s\| \leq \sqrt{\frac{D}{a_1 - \phi\left(\|\iota\|\right)}} \tag{49}$$

and can be arbitrarily small by proper control gain schedule. Finally, according to (3), we can conclude that both $e(t)$ and $\dot{e}(t)$ are bounded. More specifically, we have

$$\|e(t)\| \leq e^{-\sigma_1 t}e(0) + \frac{1}{\sigma_1}\sqrt{\frac{D}{a_1 - \phi\left(\|\iota\|\right)}}\left(1 - e^{-\sigma_1 t}\right). \tag{50}$$

References

[1] Swaroop, D., Hedrick, J., Yip, P., and Gerdes, J., 2000. "Dynamic surface control for a class of nonlinear systems". *IEEE Transactions on Automatic Control,* **45**(10), pp. 1893–1899.