

Supplementary A: proof of Theorems

Here we give the proofs for Lemma 1 and Theorems 1 and 2. We also introduce Corollary 1, which is the global \mathcal{H}_∞ version of Lemma 1.

Proof of Lemma 1

Proof of Lemma 1. We first show that (10a) - (10b) is a sufficient condition for global stability when γ_π is the global Lipschitz constant. From the assumption, we have

$$\begin{aligned} \|\delta\|_{\ell_\infty} &\leq \gamma_\Delta \|\alpha\|_{\ell_\infty} \\ &= \gamma_\Delta \|\Phi_{\alpha u} u + \Phi_{\alpha w} w + \Phi_{\alpha \delta} \delta\|_{\ell_\infty} \\ &\leq \gamma_\Delta \|\Phi_{\alpha \delta}\|_{\mathcal{L}_1} \|\delta\|_{\ell_\infty} + \gamma_\Delta \|\Phi_{\alpha u} u + \Phi_{\alpha w} w\|_{\ell_\infty}. \end{aligned}$$

From condition (10a), we can bound the ℓ_∞ norm of δ as

$$\|\delta\|_{\ell_\infty} \leq \frac{\gamma_\Delta}{1 - \beta_1} \|\Phi_{\alpha u} u + \Phi_{\alpha w} w\|_{\ell_\infty}. \quad (16)$$

Then, we have

$$\begin{aligned} \|\mathbf{y}\|_{\ell_\infty} &= \|\Phi_{\mathbf{y} \delta} \delta + \Phi_{\mathbf{y} u} u + \Phi_{\mathbf{y} w} w\|_{\ell_\infty} \quad (17a) \\ &\leq \|\Phi_{\mathbf{y} \delta}\|_{\mathcal{L}_1} \|\delta\|_{\ell_\infty} + \|\Phi_{\mathbf{y} u} u + \Phi_{\mathbf{y} w} w\|_{\ell_\infty} \quad (17b) \end{aligned}$$

$$\begin{aligned} &\leq \frac{\gamma_\Delta}{1 - \beta_1} \|\Phi_{\mathbf{y} \delta}\|_{\mathcal{L}_1} \|\Phi_{\alpha u} u + \Phi_{\alpha w} w\|_{\ell_\infty} \\ &\quad + \|\Phi_{\mathbf{y} u} u + \Phi_{\mathbf{y} w} w\|_{\ell_\infty} \quad (17c) \end{aligned}$$

$$\begin{aligned} &\leq \left(\|\Phi_{\mathbf{y} u}\|_{\mathcal{L}_1} + \frac{\gamma_\Delta \|\Phi_{\mathbf{y} \delta}\|_{\mathcal{L}_1} \|\Phi_{\alpha u}\|_{\mathcal{L}_1}}{1 - \beta_1} \right) \|u\|_{\ell_\infty} \\ &\quad + \|\Phi_{\mathbf{y} w} w\|_{\ell_\infty} + \frac{\gamma_\Delta}{1 - \beta_1} \|\Phi_{\mathbf{y} \delta}\|_{\mathcal{L}_1} \|\Phi_{\alpha w} w\|_{\ell_\infty} \quad (17d) \end{aligned}$$

$$\begin{aligned} &\leq \gamma_\pi \left(\|\Phi_{\mathbf{y} u}\|_{\mathcal{L}_1} + \frac{\gamma_\Delta \|\Phi_{\mathbf{y} \delta}\|_{\mathcal{L}_1} \|\Phi_{\alpha u}\|_{\mathcal{L}_1}}{1 - \beta_1} \right) \|y\|_{\ell_\infty} \\ &\quad + \left(\|\Phi_{\mathbf{y} w}\|_{\mathcal{L}_1} + \frac{\gamma_\Delta \|\Phi_{\mathbf{y} \delta}\|_{\mathcal{L}_1} \|\Phi_{\alpha w}\|_{\mathcal{L}_1}}{1 - \beta_1} \right) \|w\|_{\ell_\infty} \quad (17e) \end{aligned}$$

From condition (10b), we can bound the ℓ_∞ norm of \mathbf{y} as

$$\begin{aligned} \|\mathbf{y}\|_{\ell_\infty} &\leq \frac{1}{1 - \beta_2} \left(\|\Phi_{\mathbf{y} w}\|_{\mathcal{L}_1} + \frac{\gamma_\Delta \|\Phi_{\mathbf{y} \delta}\|_{\mathcal{L}_1} \|\Phi_{\alpha w}\|_{\mathcal{L}_1}}{1 - \beta_1} \right) \|w\|_{\ell_\infty}. \quad (18) \end{aligned}$$

This shows that the ℓ_∞ to ℓ_∞ gain from \mathbf{w} to \mathbf{y} is bounded. We can use similar procedure to show that the ℓ_∞ to ℓ_∞ gain from \mathbf{w} to \mathbf{u} , δ , α are all bounded. This shows the input-output stability of the closed loop system when the global Lipschitz constant of the neural network policy is γ_π .

Next, we consider the case where γ_π is valid only over a local region $\|y\|_{\ell_\infty} \leq y_\infty$. From (10c), we know that the right-hand-side of (18) is less than or equal to y_∞ given $\|w\|_{\ell_\infty} \leq w_\infty$. Thus, \mathbf{y} will never go outside the local region $\|y\|_{\ell_\infty} \leq y_\infty$ where we calculate the Lipschitz constant γ_π for any valid perturbation (7). This completes the proof. \square

The global \mathcal{H}_∞ version of Lemma 1

The following Corollary is the global \mathcal{H}_∞ version of Lemma 1.

Corollary 1. *Consider a stable LTI plant (1) - (4) interconnected with a neural network policy (5) and a dynamic uncertainty block (6) as shown in Figure 1. Assume that the persistent perturbation \mathbf{w} lies in the set given by (7). Suppose that the neural network policy $u = \pi(y)$ has a finite ℓ_2 to ℓ_2 gain γ_π for all y , and the uncertainty block Δ has the property $\|\delta\|_{\ell_2} \leq \gamma_\Delta \|\alpha\|_{\ell_2}$. If the following conditions hold:*

$$\beta_1 = \gamma_\Delta \|\Phi_{\alpha \delta}\|_{\mathcal{H}_\infty} < 1 \quad (19a)$$

$$\beta_2 = \gamma_\pi \left[\|\Phi_{\mathbf{y} u}\|_{\mathcal{H}_\infty} + \frac{\gamma_\Delta}{1 - \beta_1} \|\Phi_{\mathbf{y} \delta}\|_{\mathcal{H}_\infty} \|\Phi_{\alpha u}\|_{\mathcal{H}_\infty} \right] < 1 \quad (19b)$$

then the closed loop system in Figure 1 is ℓ_2 to ℓ_2 input-output stable.

Proof of Theorem 1

The proof of Theorem 1 relies on the concept of a positively invariant set, which is defined as follows:

Definition 1 (from (Khalil, 2002)). *A set \mathcal{M} is said to be a positively invariant set with respect to the dynamics $x[t+1] = f(x[t])$ if*

$$x[0] \in \mathcal{M} \implies x[t] \in \mathcal{M}, \forall t \geq 0. \quad (20)$$

The proof of Theorem 1 is given as follows:

Proof of Theorem 1. We use mathematical induction to show that $\mathcal{I} = \{(y, u, \alpha, \delta) \mid |y| \preceq \bar{y}, |u| \preceq \bar{u}, |\alpha| \preceq \bar{\alpha}, |\delta| \preceq \bar{\delta}\}$ is a positively invariant set of the closed loop dynamical system if the three conditions of the theorem are given. From the zero initial condition assumption in (4), we have

$$|y[0]| = |D_w w[0]| \preceq \text{abs}(D_w) \bar{w} \preceq \text{abs}(\Phi_{\mathbf{y} w}) \bar{w} \preceq \bar{y}$$

where the last inequality is from the third condition of the theorem. Then from the first condition of the theorem, we

have $|u[0]| \leq \bar{u}$. Similarly, we have

$$\begin{aligned} |\alpha[0]| &= |D_{\alpha u}u[0] + D_{\alpha w}w[0]| \\ &\leq \text{abs}(D_{\alpha u})\bar{u} + \text{abs}(D_{\alpha w})\bar{w} \\ &\leq \text{abs}(\Phi_{\alpha u})\bar{u} + \text{abs}(\Phi_{\alpha w})\bar{w} \\ &\leq \bar{\alpha}. \end{aligned}$$

Then from the second condition of the theorem, we have $|\delta[0]| \leq \bar{\delta}$. This shows $(y[0], u[0], \alpha[0], \delta[0]) \in \mathcal{I}$. Assume that we have $(y[t], u[t], \alpha[t], \delta[t]) \in \mathcal{I}$ for all $0 \leq t < T$. From the second row of the matrix equation (9), we have

$$\begin{aligned} |y[T]| &= \left| \sum_{\tau=0}^{\infty} \Phi_{yu}[\tau]u[T-\tau] + \Phi_{yw}[\tau]w[T-\tau] \right. \\ &\quad \left. + \Phi_{y\delta}[\tau]\delta[T-\tau] \right| \\ &\leq \sum_{\tau=0}^{\infty} |\Phi_{yu}[\tau]u[T-\tau]| + |\Phi_{yw}[\tau]w[T-\tau]| \\ &\quad + |\Phi_{y\delta}[\tau]\delta[T-\tau]| \\ &\leq \sum_{\tau=0}^{\infty} |\Phi_{yu}[\tau]|\bar{u} + |\Phi_{yw}[\tau]|\bar{w} + |\Phi_{y\delta}[\tau]|\bar{\delta} \\ &= \text{abs}(\Phi_{yu})\bar{u} + \text{abs}(\Phi_{yw})\bar{w} + \text{abs}(\Phi_{y\delta})\bar{\delta} \\ &\leq \bar{y} \end{aligned}$$

where the last inequality is from the third condition of the theorem. Similarly, we can derive $|\alpha[T]| \leq \bar{\alpha}$ from the third row of the matrix equation (9). The first two conditions of Theorem 1 then imply $|u[T]| \leq \bar{u}$ and $|\delta[T]| \leq \bar{\delta}$, and thus we have $(y[T], u[T], \alpha[T], \delta[T]) \in \mathcal{I}$. Using mathematical induction, we conclude that $(y[t], u[t], \alpha[t], \delta[t]) \in \mathcal{I}$ for all $t \geq 0$ and the closed loop feedback signals and state are bounded within the set specified by the theorem. \square

Proof of Theorem 2

Proof of Theorem 2. We need to show that when (10a) - (10c) and the locally Lipschitz continuous assumption of Lemma 1 are satisfied, we can always construct a quadruplet $(\bar{y}, \bar{u}, \bar{\alpha}, \bar{\delta})$ satisfying the three conditions of Theorem 1. Consider the following equation:

$$\begin{aligned} \begin{bmatrix} y_{ref} \\ \alpha_{ref} \end{bmatrix} &= \begin{bmatrix} 1 - \gamma_{\Delta}\|\Phi_{\alpha\delta}\|_{\mathcal{L}_1} & \gamma_{\Delta}\|\Phi_{y\delta}\|_{\mathcal{L}_1} \\ \gamma_{\pi}\|\Phi_{\alpha u}\|_{\mathcal{L}_1} & 1 - \gamma_{\pi}\|\Phi_{yu}\|_{\mathcal{L}_1} \end{bmatrix} \times \\ &\quad \begin{bmatrix} \|\Phi_{yw}\|_{\mathcal{L}_1} \\ \|\Phi_{\alpha w}\|_{\mathcal{L}_1} \end{bmatrix} \frac{w_{\infty}}{(1 - \beta_1)(1 - \beta_2)}, \end{aligned} \quad (23)$$

with scalar variables y_{ref} and α_{ref} . For any $w_{\infty} > 0$, we have $y_{ref} > 0$ and $\alpha_{ref} > 0$ because all the elements in (23) are positive according to the conditions (10a) - (10b). In addition, we can show that y_{ref} defined in (23) is equivalent to the left-hand-side of (10c). Therefore, we have $y_{ref} \leq y_{\infty}$ from (10c) – this means that y_{ref} is always contained within the region where we calculate the local

Lipschitz constant of the neural network policy γ_{π} . It is then straightforward to verify that $\bar{y} = y_{ref}\mathbf{1}$, $\bar{u} = \gamma_{\pi}y_{ref}\mathbf{1}$, $\bar{\alpha} = \alpha_{ref}\mathbf{1}$, $\bar{\delta} = \gamma_{\Delta}\alpha_{ref}\mathbf{1}$ satisfy the first two conditions of Theorem 1 given the locally Lipschitz continuous assumption of Lemma 1.

For the third condition of Theorem 1, it can be verified that (23) is a solution to the following inequality:

$$\begin{aligned} &\begin{bmatrix} \|\Phi_{yw}\|_{\mathcal{L}_1} \\ \|\Phi_{\alpha w}\|_{\mathcal{L}_1} \end{bmatrix} w_{\infty} \\ &\leq \begin{bmatrix} 1 - \gamma_{\pi}\|\Phi_{yu}\|_{\mathcal{L}_1} & -\gamma_{\Delta}\|\Phi_{y\delta}\|_{\mathcal{L}_1} \\ -\gamma_{\pi}\|\Phi_{\alpha u}\|_{\mathcal{L}_1} & 1 - \gamma_{\Delta}\|\Phi_{\alpha\delta}\|_{\mathcal{L}_1} \end{bmatrix} \begin{bmatrix} y_{ref} \\ \alpha_{ref} \end{bmatrix}, \end{aligned}$$

which can be rearranged into

$$\begin{aligned} &\begin{bmatrix} \|\Phi_{yw}\|_{\mathcal{L}_1} \\ \|\Phi_{\alpha w}\|_{\mathcal{L}_1} \end{bmatrix} w_{\infty} + \begin{bmatrix} \|\Phi_{yu}\|_{\mathcal{L}_1} & \|\Phi_{y\delta}\|_{\mathcal{L}_1} \\ \|\Phi_{\alpha u}\|_{\mathcal{L}_1} & \|\Phi_{\alpha\delta}\|_{\mathcal{L}_1} \end{bmatrix} \begin{bmatrix} \gamma_{\pi}y_{ref} \\ \gamma_{\Delta}\alpha_{ref} \end{bmatrix} \\ &\leq \begin{bmatrix} y_{ref} \\ \alpha_{ref} \end{bmatrix}. \end{aligned} \quad (24)$$

Finally, we note the following inequality

$$\text{abs}(\Phi)\mathbf{1} \leq \|\Phi\|_{\mathcal{L}_1}\mathbf{1} \quad (25)$$

from the fact that the \mathcal{L}_1 norm is selecting the maximum row sum. Therefore, we have

$$\begin{aligned} &\text{abs} \left(\begin{bmatrix} \Phi_{yw} \\ \Phi_{\alpha w} \end{bmatrix} \right) \bar{w} + \text{abs} \left(\begin{bmatrix} \Phi_{yu} & \Phi_{y\delta} \\ \Phi_{\alpha u} & \Phi_{\alpha\delta} \end{bmatrix} \right) \begin{bmatrix} \bar{u} \\ \bar{\delta} \end{bmatrix} \\ &\leq \begin{bmatrix} \|\Phi_{yw}\|_{\mathcal{L}_1}\mathbf{1} \\ \|\Phi_{\alpha w}\|_{\mathcal{L}_1}\mathbf{1} \end{bmatrix} w_{\infty} + \text{abs} \left(\begin{bmatrix} \Phi_{yu} & \Phi_{y\delta} \\ \Phi_{\alpha u} & \Phi_{\alpha\delta} \end{bmatrix} \right) \begin{bmatrix} \gamma_{\pi}y_{ref}\mathbf{1} \\ \gamma_{\Delta}\alpha_{ref}\mathbf{1} \end{bmatrix} \\ &\leq \begin{bmatrix} \|\Phi_{yw}\|_{\mathcal{L}_1}\mathbf{1} \\ \|\Phi_{\alpha w}\|_{\mathcal{L}_1}\mathbf{1} \end{bmatrix} w_{\infty} \\ &\quad + \begin{bmatrix} \|\Phi_{yu}\|_{\mathcal{L}_1}\mathbf{1} & \|\Phi_{y\delta}\|_{\mathcal{L}_1}\mathbf{1} \\ \|\Phi_{\alpha u}\|_{\mathcal{L}_1}\mathbf{1} & \|\Phi_{\alpha\delta}\|_{\mathcal{L}_1}\mathbf{1} \end{bmatrix} \begin{bmatrix} \gamma_{\pi}y_{ref} \\ \gamma_{\Delta}\alpha_{ref} \end{bmatrix} \\ &\leq \begin{bmatrix} y_{ref}\mathbf{1} \\ \alpha_{ref}\mathbf{1} \end{bmatrix} \\ &= \begin{bmatrix} \bar{y} \\ \bar{\alpha} \end{bmatrix}. \end{aligned}$$

We can see that if the conditions of Lemma 1 hold, we can always construct a quadruplet $(\bar{y}, \bar{u}, \bar{\alpha}, \bar{\delta})$ satisfying all the three conditions of Theorem 1. The converse is not true. Therefore, Theorem 1 can be applied on a strictly larger class of problems than Lemma 1. \square

Supplementary B: cart pole model

We consider a cart-pole problem with η being the displacement of the cart and θ the angle of the pole. The dynamics

is given by

$$\begin{aligned}\ddot{\eta} &= \left(\frac{4}{3}(M+m)l - ml\cos^2(\theta) \right)^{-1} \\ &\quad \left(\frac{4}{3}ml^2\dot{\theta}^2\sin(\theta) - mgl\sin(\theta)\cos(\theta) + \frac{4}{3}lu \right) \\ \ddot{\theta} &= \left(\frac{4}{3}(M+m)l - ml\cos^2(\theta) \right)^{-1} \\ &\quad \left(-ml\dot{\theta}^2\sin(\theta)\cos(\theta) + (M+m)g\sin(\theta) - \cos(\theta)u \right).\end{aligned}$$

We use the default model parameters from stable baselines: $g = 9.8$, $M = 1$, $m = 0.1$, and $l = 0.5$. Using Euler discretization, the nonlinear discrete time cart pole model is given by

$$\begin{aligned}\eta[t+1] &= \eta[t] + \tau\dot{\eta}[t] \\ \dot{\eta}[t+1] &= \dot{\eta}[t] + \tau\ddot{\eta}[t] \\ \theta[t+1] &= \theta[t] + \tau\dot{\theta}[t] \\ \dot{\theta}[t+1] &= \dot{\theta}[t] + \tau\ddot{\theta}[t]\end{aligned}$$

with sampling time $\tau = 0.02$.

Linearized model with no uncertainty

Define the state vector $x = [\eta \quad \dot{\eta} \quad \theta \quad \dot{\theta}]^\top$. The linearized cart-pole model around the origin is given by

$$x[t+1] = Ax[t] + Bu[t], \quad y[t] = x[t] + D_w w[t]$$

with

$$\begin{aligned}A &= \begin{bmatrix} 1 & \tau & 0 & 0 \\ 0 & 1 & \frac{-3mg\tau}{4M+m} & 0 \\ 0 & 0 & 1 & \tau \\ 0 & 0 & \frac{3(M+m)g\tau}{(4M+m)l} & 1 \end{bmatrix}, \\ B &= \begin{bmatrix} 0 \\ \frac{4\tau}{4M+m} \\ 0 \\ \frac{-3\tau}{(4M+m)l} \end{bmatrix}, \quad D_w = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.\end{aligned}\quad (28)$$

Note from D_w that we have an one dimensional perturbation on the pole angle measurement. The requirement is to certify that the actual pole angle is within the user-specified limit, that is,

$$|x_3[t]| \leq x_{lim}, \quad \forall t \geq 0. \quad (29)$$

As explained in the Case Study Section, the assumptions of one dimensional perturbation and one dimensional state constraint are made for the ease of illustrating the result. Algorithm 1 can be applied to systems with multi-dimensional perturbations with user-specified requirement on both state, measurement, and control action.

Learned model using data driven approach

When the model equation is unknown in the first place, we can collect the data from the simulator and fit the data to a model. Let $x^{(i,j)}[t]$ be the state vector x at time t for the i -th episode from the j -th bootstrap sample, for $j = 0, 1, \dots, 100$. For the j -th bootstrap run, we solve the following least square problem to obtain the system matrices $A^{(j)}$ and $B^{(j)}$:

$$\begin{aligned}\text{minimize}_{A^{(j)}, B^{(j)}} & \sum_{i=1}^N \sum_{t=0}^{T-1} \|x^{(i,j)}[t+1] - A^{(j)}x^{(i,j)}[t] \\ & - B^{(j)}u^{(i,j)}[t]\|_2^2\end{aligned}\quad (30)$$

with $T = 30$, x from the simulator, and u randomly generated. We then find a pair of non-negative matrices Δ_A and Δ_B such that

$$|A^{(j)} - A^{(0)}| \preceq \Delta_A, \quad |B^{(j)} - B^{(0)}| \preceq \Delta_B, \quad (31)$$

for $j = 1, \dots, 100$ to over-approximate the modeling error of the nominal model $(A^{(0)}, B^{(0)})$. The learned model used by our experiment in the Case Study Section is then given by

$$\begin{aligned}x[t+1] &= A^{(0)}x[t] + B^{(0)}u[t] + \delta[t] \\ y[t] &= x[t] + D_w w[t] \\ \alpha[t] &= \begin{bmatrix} I \\ 0 \end{bmatrix} x[t] + \begin{bmatrix} 0 \\ I \end{bmatrix} u[t] = \begin{bmatrix} x[t] \\ u[t] \end{bmatrix} \\ |\delta[t]| &\preceq \begin{bmatrix} \Delta_A & \Delta_B \end{bmatrix} |\alpha[t]| = \Gamma_\Delta |\alpha[t]|,\end{aligned}$$

where $\Gamma_\Delta = \begin{bmatrix} \Delta_A & \Delta_B \end{bmatrix}$ is the non-negative matrix used in Algorithm 1.