Automatica 146 (2022) 110622

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Dynamic state feedback controller and observer design for dynamic artificial neural network models ☆

Anastasia Nikolakopoulou [1], Moo Sun Hong [1], Richard D. Braatz *,[1]

Department of Chemical Engineering, Massachusetts Institute of Technology, 02139, Cambridge, MA, USA

ARTICLE INFO

Article history:
Received 7 March 2021
Received in revised form 15 May 2022
Accepted 10 August 2022
Available online xxxx

Keywords:
Dynamic state feedback
State observer
Output observer
Neural network controller
Linear matrix inequalities

ABSTRACT

Artificial neural networks are black-box models that can be used to model nonlinear dynamical systems. This article presents a synthesis method for full dynamic state feedback controllers and state and output observers that have guaranteed properties for systems approximated by dynamic artificial neural networks. The resulting control designs are applicable to the practical situation in which the steady-state values for the control input are not known. Dynamic artificial neural networks are written in the standard nonlinear operator form, also known in the literature as the Luré formulation. A generalized form of the Luré formulation is adopted to allow for the representation of deep ℓ-layer networks, $\ell \geq 1$. Sufficient conditions for controller synthesis and observer design are derived in the form of linear matrix inequalities, using a quadratic Lyapunov function. The synthesis method is demonstrated for the control of pH in two tanks in series and a numerical example.

1. Introduction

Artificial neural networks (ANNs) have been used in nonlinear system identification and control for decades (Himmelblau, 2008; Pearson, 1995) due to their ability to universally approximate any static nonlinear function (Cybenko, 1989; Funahashi, 1989). Dynamic model structures that incorporate ANNs, known as *dynamic artificial neural networks* (DANNs), can universally approximate any nonlinear dynamic relationship (e.g., see Kim, Ríos-Patrón, & Braatz, 2018 and citations therein). This universality property holds for many DANN model structures, including neural state space models (NSSMs) (Suykens, De Moor, & Vandewalle, 1995), global input–output models (GIOMs) (Levin & Narendra, 1995), and dynamic recurrent neural networks (DRNNs) (Hopfield, 1982; Michel, Farrell, & Porod, 1989).

Stability and performance analysis for open- and closed-loop DANNs has been of long-term interest. DANNs can be represented as systems with linear time-invariant dynamics in feedback with a static nonlinear operator that is diagonal, continuous, and sector bounded (Ríos-Patrón & Braatz, 1998). Block diagram algebra can be used to represent DANNs in standard nonlinear operator form (SNOF), for which sufficient stability criteria can exploit properties of the activation functions such as sector boundedness and slope restriction (Kim et al., 2018). This SNOF representation is closely related to Luré systems (Luré & Postnikov, 1944), and is general enough to be applicable to deep neural networks. The SNOF is the underlying structure in several neural network-based analyses, such as a recent work on safety verification and robustness of neural networks (Fazlyab, Morari, & Pappas, 2020) and stability certification of deep learning-based controllers (Nguyen et al., 2020).

The literature on stability analysis for Luré systems is vast. A standard method to examine the absolute stability of a Luré system is the search for the existence of appropriate Lyapunov functions. The quadratic and the Luré–Postnikov functions are popular Lyapunov function candidates (Khalil, 2002). Alternative formulations include the modified Luré–Postnikov functions which are quadratic in the states and the nonlinearity and include an additional integral term (Drummond, Valmorbida, & Duncan, 2018; Park, 2002), as well as a Lyapunov function which includes cross terms between the state and the cone bounded nonlinearity (Gonzaga, Jungers, & Daafouz, 2012). Luré–Postnikov and their modified counterpart function candidates may reduce conservatism by taking properties of the nonlinearity into account, such as slope restrictions. The reduced conservatism comes at the expense of more complicated matrix inequality formulations unless simplifications in the system formulation or the Lyapunov function are introduced.

☆ This work was supported by the DARPA Make-It program under contract ARO W911NF-16-2-0023. The material in this paper was partially presented at the 59th IEEE Conference on Decision and Control, December 14–18, 2020, Jeju Island, Republic of Korea and the 2021 American Control Conference, May 25–28, 2021, New Orleans, LA, USA. This paper was recommended for publication in revised form by Associate Editor Angelo Alessandri under the direction of Editor Thomas Parisini.

* Corresponding author.
E-mail addresses: anikol@mit.edu (A. Nikolakopoulou), moosun1@mit.edu (M.S. Hong), braatz@mit.edu (R.D. Braatz).

[1] Authors contributed equally to the paper.

https://doi.org/10.1016/j.automatica.2022.110622
0005-1098/

A common simplification in the Luré literature is to omit the term that allows for the nonlinearity to be an input to its own argument ($D_{qp} = 0$ in $q = C_q x + D_{qp} p$, where $p = \phi(q)$) (Gonzaga et al., 2012; Park, 2002). This omission, while greatly simplifying the derivations of stability proofs, prevents the existing work from being applicable to ℓ-layer neural networks, $\ell > 1$. When solving the controller synthesis problem for Luré systems, higher order matrix inequalities arise, which requires special handling in order to obtain linear matrix inequality (LMI) design criteria. A quadratic Lyapunov function is often adopted for the synthesis problem (Diwadkar, Dasgupta, & Vaidya, 2015; Kim & Braatz, 2014; Suykens, Curran, & Chua, 1999), in order to facilitate the handling of the higher order matrix inequalities. Under the assumption of $D_{qp} = 0$ for discrete-time Luré systems, design results have been derived for state observers and dynamic output feedback controllers (Kim & Braatz, 2014) and observer-based robust feedback controllers (Diwadkar et al., 2015).

We have presented sufficient criteria for dynamic state feedback controller synthesis and state observer design for continuous-time DANNs reformulated as Luré systems for the important case where $D_{qp} \neq 0$ (Nikolakopoulou, Hong, & Braatz, 2020). Additionally, we have derived criteria for dynamic output feedback controller synthesis and output observers for discrete-time DANNs (Nikolakopoulou, Hong, & Braatz, 2021). These works were demonstrated on case studies for single-node neural network models. This article is a major extension of the aforementioned conference papers (Nikolakopoulou et al., 2020, 2021). Based on the discrete-time Luré reformulation, we derive LMI criteria for full dynamic state feedback controller synthesis. Moreover, we derive LMI criteria for nonlinear full dynamic state feedback controller design. Additionally, we present an alternative formulation for the observer design than used in the conference articles (Nikolakopoulou et al., 2020, 2021) and derive LMI design criteria. The approaches are demonstrated for the design of observers and control systems in two numerical examples.

This article is structured as follows. Section 2 presents some preliminaries. Section 3 introduces alternative representations for DANNs, including deep networks, and we show that the $D_{qp} = 0$ assumption made in literature studies to simplify the theoretical analysis also greatly restricts the applicability of such results for analyzing DANNs. Section 4 discusses a set of LMIs for obtaining stabilizing dynamic state feedback controllers for DANNs formulated in SNOF and as Luré systems sector bounded in $[-1, 1]$. Section 5 derives LMI criteria for designing state observers that maximize the decay rate of the observer error dynamics. LMI criteria to obtain output observers with guaranteed L_2-gain (from an unmeasured disturbance to the process output) are presented in Section 6. The theoretical results are illustrated in two case studies in Section 7. Lastly, conclusions are summarized in Section 8.

2. Background

2.1. Notation

The superscript $^\top$ stands for the matrix transpose. $\mathbb{Z}_{0+}$ is the set of non-negative integers, $\mathbb{R}^n$ denotes the n-dimensional Euclidean space, $e_i \in \mathbb{R}^n$ is the ith unit vector, and $\mathbb{R}^{n\times m}$ is the set of all $n \times m$ real matrices. Positive and negative definiteness of a matrix are denoted by $\succ$ and $\prec$ respectively, and positive and negative semidefiniteness by $\succeq$ and $\preceq$ respectively. The set of symmetric $n \times n$ matrices is denoted by $\mathbb{S}^n = \{X \in \mathbb{R}^{n\times n} \mid X = X^\top\}$. The notation $\mathbb{S}^n_{++}$ is used to denote the set of symmetric positive definite matrices $\mathbb{S}^n_{++} = \{X \in \mathbb{S}^n \mid X \succ 0\}$. The set of diagonal $n\times n$ matrices with elements x_i is denoted by $\mathbb{D}^n = \{X \in \mathbb{R}^{n\times n} \mid X = \mathrm{diag}\{x_i\},\ i = 1, \dots, n\}$. The notations $\mathbb{D}^n_{++}$ and $\mathbb{D}^n_{+}$ are used to denote the set of diagonal positive definite matrices $\mathbb{D}^n_{++} = \{X \in \mathbb{D}^n \mid X \succ 0\}$ and diagonal positive semidefinite matrices $\mathbb{D}^n_{+} = \{X \in \mathbb{D}^n \mid X \succeq 0\}$ respectively. The variable x at the kth time instance is described by x_k. Corresponding symmetric matrix elements are replaced by $*$. The Euclidean norm of vector $x \in \mathbb{R}^n$ is denoted by $\|x\|$. For a nonlinear function applied element-wise to the vector $q \in \mathbb{R}^{n_q}$ such that $\varphi(q) = [\phi_1(q_1) \cdots \phi_{n_q}(q_{n_q})]^\top$, define

- Sector boundedness: $\Phi^{[\alpha,\beta]}_{sb}$ is used to describe the set

$$\{\varphi : \mathbb{R}^{n_q} \to \mathbb{R}^{n_q} \mid [\phi_i(\sigma) - \alpha_i\sigma][\phi_i(\sigma) - \beta_i\sigma] \le 0, \quad \forall\sigma \in \mathbb{R}, i = 1, \dots, n_q\}.$$

- Slope restriction: $\Phi^{[\mu_1,\mu_2]}_{sr}$ is used to describe the set

$$\{\varphi : \mathbb{R}^{n_q} \to \mathbb{R}^{n_q} \mid \mu_{1,i} \le \tfrac{\phi_i(\sigma_1)-\phi_i(\sigma_2)}{\sigma_1-\sigma_2} \le \mu_{2,i}, \quad \forall\sigma_1 \neq \sigma_2 \in \mathbb{R}, i = 1, \dots, n_q\}.$$

Definition 1 (*Stability in the Sense of Lyapunov*)**.** Given a nonlinear time-varying discrete-time system of the form

$$x(k+1) = f(k, x(k)),\ \ x(k_0) = x_0, \tag{1}$$

where $x \in \mathbb{R}^{n_x}$, $k \in \mathbb{Z}_{0+}$, the equilibrium state 0 of (1) is stable in the sense of Lyapunov if, for every $\epsilon > 0$, there exists a $\delta(\epsilon, k_0) > 0$ such that, if $\|x_0\| < \delta$, then $\|x(k)\| < \epsilon$ for all $k > k_0$.

Definition 2 (*Local Asymptotic Stability*)**.** The equilibrium state 0 of (1) is locally asymptotically stable if (a) the state 0 is stable in the sense of Lyapunov and (b) there exists a $\delta'(k_0)$ such that, if $\|x_0\| < \delta'$, then $x(k) \to 0$ as $k \to \infty$.

Definition 3 (*Global Asymptotic Stability*)**.** The equilibrium state 0 of (1) is globally asymptotically stable (g.a.s.) if the state is asymptotically stable for any $\delta' > 0$.

2.2. Matrix inequalities

A strict linear matrix inequality has the form

$$F(x) \triangleq F_0 + \textstyle\sum_{i=1}^m x_i F_i \succ 0, \tag{2}$$

where $x \in \mathbb{R}^m$ is the variable and the matrices $F_i \in \mathbb{S}^n$, $i = 1, \dots, m$ are known. For a nonstrict LMI, the condition that $F(x) \succ 0$ is replaced by $F(x) \succeq 0$. LMIs are convex, and formulating optimizations with convex objective functions subject to LMI constraints is appealing due to the existence of efficient algorithms (Bland, Goldfarb, & Todd, 1981; Nesterov & Nemirovskii, 1994) and readily available software that can solve such problems, e.g. Chiang (2019). Common techniques used in LMI methodologies for the derivation of performance and controller synthesis criteria, such as the Schur-complement lemma and the S-procedure, are used in the proofs. Readers who would like more details on these techniques are referred to Boyd, El Ghaoui, Feron, and Balakrishnan (1994).

2.3. S-procedure

The S-procedure for quadratic forms (Boyd et al., 1994; Gusev & Likhtarnikov, 2006) for the strict case states that, if there exist $\tau_1 \ge 0, \dots, \tau_p \ge 0$ such that $T_0 - \sum_{i=1}^p \tau_i T_i \succ 0$, then $x^\top T_0 x > 0$, $\forall x \neq 0 \ni x^\top T_i x \ge 0$, $i = 1, \dots, p$, where $T_0, T_i \in \mathbb{R}^{n\times n}$. For the non-strict case, the S-procedure indicates that if there exist $\tau_1 \ge 0, \dots, \tau_p \ge 0$ such that $T_0 - \sum_{i=1}^p \tau_i T_i \succeq 0$, then $x^\top T_0 x \ge 0$, $\forall x \ni x^\top T_i x \ge 0$, $i = 1, \dots, p$.

2.4. Luré system

The Luré system for the discrete-time case is often encountered in the form (Diwadkar et al., 2015; Gonzaga et al., 2012; Kim & Braatz, 2014; Park, 2002; Suykens et al., 1999)

$$\begin{aligned} x_{k+1} &= Ax_k + B_p p_k, \\ q_k &= C_q x_k, \\ p_{k,i} &= \phi_i(q_{k,i}),\ i = 1,\ldots,n_q, \end{aligned} \tag{3}$$

where subscript k refers to the kth time instance, $x_k \in \mathbb{R}^{n_x}$ is the vector of system states, $p_k \in \mathbb{R}^{n_q}$ is the vector of the nonlinearities, $q_k \in \mathbb{R}^{n_q}$ is the argument vector of the nonlinear operator, ϕ_i is the ith diagonal element of the nonlinear operator $\varphi \in \Phi_{sb}^{[\xi_1,\xi_2]}$, $\xi_2 > \xi_1$, with $\phi_i(0) = 0$. Eq. (3) is in the form of a recurrent neural network with one hidden layer, n_q nodes, activation functions $\phi_i(\cdot)$, and matrices $A \in \mathbb{R}^{n_x \times n_x}$, $B_p \in \mathbb{R}^{n_x \times n_q}$, $C_q \in \mathbb{R}^{n_q \times n_x}$.

2.5. Problem statement

The Luré system (3) only describes neural networks with one hidden layer. This article considers analysis and observer and controller design for more general system descriptions (6) and (12) which are applicable to neural networks with multiple hidden layers.

3. Alternative representations of dynamic artificial neural networks

In systems applications, neural networks with more than one layer may be employed for predictive modeling. Multilayer neural networks can be written directly as a SNOF. Here we demonstrate the reformulation for the general ℓ-layer recurrent deep network

$$\begin{aligned} p_k^1 &= \phi_1(W_x^0 x_k + W_u^0 u_k + b^0), \\ p_k^{j+1} &= \phi_{j+1}(W^j p_k^j + b^j),\ \ j = 1,\ldots,\ell-1, \\ x_{k+1} &= Ax_k + B_u u_k + W^\ell p_k^\ell + b^\ell, \end{aligned} \tag{4}$$

or equivalently

$$\begin{aligned} p_k &= [p_k^1\ p_k^2\ \ldots\ p_k^\ell]^\top = \varphi(q_k), \\ q_k &= \begin{bmatrix} q_k^1 \\ q_k^2 \\ \vdots \\ q_k^\ell \end{bmatrix} = \begin{bmatrix} W_x^0 x_k + W_u^0 u_k + b^0 \\ W^1 p_k^1 + b^1 \\ \vdots \\ W^{\ell-1} p_k^{\ell-1} + b^{\ell-1} \end{bmatrix}, \\ x_{k+1} &= Ax_k + B_u u_k + W^\ell p_k^\ell + b^\ell, \end{aligned} \tag{5}$$

where $u_k \in \mathbb{R}^{n_u}$ is the exogenous input, $q_k^j \in \mathbb{R}^{n_j}$ is the pre-activation vector of the jth layer, n_j is the number of nodes in the jth layer, $W_x^0 \in \mathbb{R}^{n_1 \times n_x}$ and $W_u^0 \in \mathbb{R}^{n_1 \times n_u}$ are the weight matrices of the first layer corresponding to the neural network inputs x_k and u_k, $W^j \in \mathbb{R}^{n_{j+1} \times n_j}$ and $b^j \in \mathbb{R}^{n_{j+1}}$ are the weight matrix and bias vector of the $(j+1)$th layer, and x_{k+1} is the neural network output. The nonlinear function φ is applied element-wise to q_k.

The ℓ-layer neural network (4) is reformulated as a SNOF with a non-zero D_{qp} matrix,

$$\begin{aligned} x_{k+1} &= Ax_k + B_p p_k + B_u u_k + b_c,\quad x(0) = x_0, \\ q_k &= C_q x_k + D_{qp} p_k + D_{qu} u_k + d_{qc}, \\ p_{k,i} &= \phi_i(q_{k,i}),\ i = 1,\ldots,n_q = \textstyle\sum_{j=1}^{\ell} n_j, \end{aligned} \tag{6}$$

Table 1
Characteristics of common DANN activation functions.

Activation function	Sector bound	Slope restriction
sigmoid	[0, 0.25]	(0, 0.25]
tanh	[0, 1]	(0, 1]
ReLU	[0, 1]	[0, 1]
leaky ReLU, $a < 1$	$[a, 1]$	$[a, 1]$
exponential LU, $a \geq 0$	[0, 1]	(0, 1]
SoftPlus	[0, 1]	(0, 1)

where

$$B_p = [0\ \cdots\ 0\ W^\ell],\quad b_c = b^\ell,\quad C_q = [W_x^0\ 0\ \cdots\ 0]^\top,$$
$$D_{qp} = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ W^1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & W^{\ell-1} & 0 \end{bmatrix},\ D_{qu} = \begin{bmatrix} W_u^0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ d_{qc} = \begin{bmatrix} b^0 \\ b^1 \\ \vdots \\ b^{\ell-1} \end{bmatrix}, \tag{7}$$

and $\varphi \in \Phi_{sb}^{[\xi_1,\xi_2]}$, $\xi_2 > \xi_1$. Measured and unmeasured disturbances can be modeled as part of the exogenous input vector u_k. In the formulation (6), x_k represents the system states and not the neural network states. This representation is not limited to recurrent neural networks. For example, neural nets with backward connectivity would be described by a D_{qp} matrix that has its upper triangular area populated as well. The above analysis illustrates the value of considering the more general formulation of (3) when analyzing and designing estimators and control systems for neural networked systems. Sector boundedness and slope restriction characteristics of common DANN activation functions are given in Table 1. The sector bounds for the sigmoid and the SoftPlus functions refer to the sector bounds of the resulting function, f, after the original function has been shifted by a constant such that $f(0) = 0$.

A loop transformation with constant gains from the sector $[\xi_{1,i}, \xi_{2,i}]$ to $[-1, 1]$ for all $i = 1,\ldots,n_q$ can be used. This transformation requires well-posedness of (6), i.e. $\det(I - D_{qp}\Delta) \neq 0$ for all $\Delta = \mathrm{diag}\{\Delta_{ii}\}$, with $\xi_{1,i} \leq \Delta_{ii} \leq \xi_{2,i}$.

The transformation results in

$$\begin{aligned} x_{k+1} &= \bar{A}x_k + \bar{B}_p\bar{p}_k + \bar{B}_u u_k + \bar{b}_c,\quad x(0) = x_0, \\ q_k &= \bar{C}_q x_k + \bar{D}_{qp}\bar{p}_k + \bar{D}_{qu}u_k + \bar{d}_{qc}, \\ \bar{p}_{k,i} &= \delta_i(k)q_{k,i},\quad |\delta_i(k)| \leq 1,\quad i = 1,\ldots,n_q, \end{aligned} \tag{8}$$

where

$$\bar{p}_{k,i} \triangleq \tfrac{2}{\xi_{2,i}-\xi_{1,i}}(p_{k,i} - \xi_{1,i}q_{k,i}) - q_{k,i}. \tag{9}$$

Define

$$\Lambda \triangleq \mathrm{diag}\left\{\tfrac{\xi_{1,i}+\xi_{2,i}}{2}\right\},\ \ \tilde{\Lambda} \triangleq \mathrm{diag}\left\{\tfrac{\xi_{2,i}-\xi_{1,i}}{2}\right\}, \tag{10}$$

then the transformed matrices are (Nikolakopoulou et al., 2020)

$$\begin{aligned} \bar{A} &= A + B_p\Lambda(I - D_{qp}\Lambda)^{-1}C_q, \\ \bar{B}_p &= B_p\tilde{\Lambda} + B_p\Lambda(I - D_{qp}\Lambda)^{-1}D_{qp}\tilde{\Lambda}, \\ \bar{B}_u &= B_u + B_p\Lambda(I - D_{qp}\Lambda)^{-1}D_{qu}, \\ \bar{b}_c &= b_c + B_p\Lambda(I - D_{qp}\Lambda)^{-1}d_{qc}, \\ \bar{C}_q &= (I - D_{qp}\Lambda)^{-1}C_q,\quad \bar{D}_{qp} = (I - D_{qp}\Lambda)^{-1}D_{qp}\tilde{\Lambda}, \\ \bar{D}_{qu} &= (I - D_{qp}\Lambda)^{-1}D_{qu},\quad \bar{d}_{qc} = (I - D_{qp}\Lambda)^{-1}d_{qc}. \end{aligned} \tag{11}$$

The perturbation $\delta_i(k)$ can be used interchangeably with $|\bar{p}_{k,i}| = |\bar{\phi}_i(\sigma)| \leq |\sigma|$, both suggesting that $|\bar{p}_{k,i}| \leq |q_{k,i}|$. More details on the derivation and the matrix transformations can be found in Nikolakopoulou et al. (2020).

4. Controller synthesis

Here, we formulate the dynamic state feedback synthesis problem for discrete-time DANNs reformulated as Luré systems, sector bounded in $[-1,1]$. A static state feedback control law in deviation variables is $u-u_{ss}=K(x-x_{ss})$, with u_{ss} being the value of the control input at steady state. However, the true relationship between u_{ss} and x_{ss} is not known in practical applications, especially under the presence of model uncertainty. Incorporating an integrator into the state feedback, which results in a dynamic state feedback controller, removes the need to know this relationship.

Consider the system described by

$$\begin{aligned} x_{k+1}&=Ax_k+B_pp_k+B_uu_k+B_dd_k+b_c,\quad x(0)=x_0,\\ I_{k+1}&=I_k+C_yx_k-r_k,\quad I_0=0,\\ q_k&=C_qx_k+D_{qp}p_k+D_{qu}u_k+D_{qd}d_k+d_{qc},\\ p_k&=\varphi(q_k),\ p_{k,i}=\phi_i(q_{k,i}),\ |\phi_i(\sigma)|\le|\sigma|, \end{aligned}\tag{12}$$

where $u_k\in\mathbb{R}^{n_u}$ is the control input vector, $d_k\in\mathbb{R}^{n_d}$ is the measured disturbance vector, $I_k\in\mathbb{R}^{n_y}$ is the integrator state vector, $r_k\in\mathbb{R}^{n_y}$ is the desired output setpoint vector which can be constant or time-varying, constant at steady state. In a Lyapunov framework, we derive criteria for the stability of (12) in deviation variables. For global asymptotic stability, $\tilde{x}_k\triangleq x_k-x_{ss}\to 0$ as $k\to+\infty$. Any assumed value for x_{ss} implicitly assumes some known steady-state value for the disturbance $d=d_{ss}$ since x_{ss} changes if the steady-state disturbance changes. Additionally, define the deviation variables $\forall i=1,\dots,n_q$,

$$\begin{aligned} \tilde{p}_i&\triangleq p_i-p_{i,ss}=\phi_i(q_i)-\phi_i(q_{i,ss}):=\tilde{\phi}_i(q_i-q_{i,ss};q_{i,ss}),\\ \tilde{p}&\triangleq p-p_{ss}=\varphi(q)-\varphi(q_{ss}):=\tilde{\varphi}(q-q_{ss};q_{ss}). \end{aligned}$$

Below are some easy-to-prove lemmas used in the proofs of theorems later in this article.

Lemma 1. *For all ϕ_i that is slope restricted in $[\mu_{1,i},\mu_{2,i}]$, $\tilde{\phi}_i(\cdot;q_{i,ss})$ is sector bounded in $[\mu_{1,i},\mu_{2,i}]$.*

Lemma 2. *If $\varphi\in\Phi_{sb}^{[\xi_1,\xi_2]}$ and $\varphi\in\Phi_{sr}^{[\xi_1,\xi_2]}$, then $\tilde{\varphi}\in\Phi_{sb}^{[\xi_1,\xi_2]}$.*

Rewriting (12) in terms of deviation variables and applying Lemma 2 gives

$$\begin{aligned} \tilde{z}_{k+1}&=\begin{bmatrix}A&0\\C_y&I\end{bmatrix}\tilde{z}_k+\begin{bmatrix}B_p\\0\end{bmatrix}\tilde{p}_k+\begin{bmatrix}B_u\\0\end{bmatrix}\tilde{u}_k,\\ \tilde{q}_k&=[C_q\ 0]\tilde{z}_k+D_{qp}\tilde{p}+D_{qu}\tilde{u}_k,\\ \tilde{p}_{k,i}&=\tilde{\phi}_i(\tilde{q}_{k,i}),\ |\tilde{\phi}_i(\sigma)|\le|\sigma|, \end{aligned}\tag{13}$$

where $\tilde{z}_k=[\tilde{x}_k\ \tilde{I}_k]$ and I is the identity matrix. For a full dynamic state feedback control law $\tilde{u}_k=K_p\tilde{x}_k+K_i\tilde{I}_k=\bar{K}\tilde{z}_k$, $\bar{K}=[K_p\ K_i]$, (13) can be rewritten as

$$\begin{aligned} \tilde{z}_{k+1}&=\tilde{A}\tilde{z}_k+B'_p\tilde{p}_k,\\ \tilde{q}_k&=\tilde{C}_q\tilde{z}_k+D_{qp}\tilde{p}_k,\\ \tilde{p}_{k,i}&=\tilde{\phi}_i(\tilde{q}_{k,i}),\ |\tilde{\phi}_i(\sigma)|\le|\sigma|, \end{aligned}\tag{14}$$

where

$$\begin{aligned} &\tilde{A}\triangleq A'+B'_u\bar{K},\ \tilde{C}_q\triangleq C'_q+D_{qu}\bar{K},\\ &A'=\begin{bmatrix}A&0\\C_y&I\end{bmatrix},\ B'_u=\begin{bmatrix}B_u\\0\end{bmatrix},\ B'_p=\begin{bmatrix}B_p\\0\end{bmatrix},\ C'_q=[C_q\ 0]. \end{aligned}\tag{15}$$

4.1. Full dynamic state feedback controller for a SNOF sector bounded in $[\xi_1,\xi_2]$

This section derives criteria for stabilizing controller synthesis for systems of the structure (14) with different sector bounds:

$$\begin{aligned} \tilde{z}_{k+1}&=\tilde{A}\tilde{z}_k+B'_p\tilde{p}_k,\\ \tilde{q}_k&=\tilde{C}_q\tilde{z}_k+D_{qp}\tilde{p}_k,\\ \tilde{p}_{k,i}&=\tilde{\phi}_i(\tilde{q}_{k,i}),\ \tilde{\varphi}\in\Phi_{sb}^{[\xi_1,\xi_2]},\ \xi_2>\xi_1, \end{aligned}\tag{16}$$

where the matrices are defined as in (15). Sufficient criteria for stabilizing controller synthesis of (16) result in nonlinear matrix inequalities. Assuming less general sector bounds (or employing a loop transformation) allows for simplifications that can lead to lower order matrix inequalities (Diwadkar et al., 2015; Kim & Braatz, 2014; Suykens et al., 1999). Many common DANN activation functions (see Table 1) are sector bounded in $[0,\xi]$, $\xi>0$ (with the exception of leaky ReLU which is sector bounded in $[\xi_1,\xi_2]$, $0<\xi_1<\xi_2$). Stabilizing controller synthesis criteria for such systems could be obtained by the following theorem.

Theorem 1. *A sufficient condition for the global asymptotic stability of the controlled system (16) with $\xi_2=\xi>\xi_1=0$ is the existence of $Q\in\mathbb{S}_{++}^{n_x+n_y}$, $\tilde{M}\in\mathbb{D}_{++}^{n_q}$, and $Y\in\mathbb{R}^{n_u\times(n_x+n_y)}$ that satisfy the LMI*

$$\mathrm{SF}_{[0,\xi]}\triangleq\begin{bmatrix}-Q&*&*\\ \begin{pmatrix}\frac{1}{2}\Xi C'_qQ\\+\frac{1}{2}\Xi D_{qu}Y\end{pmatrix}&\begin{pmatrix}-\tilde{M}+\frac{1}{2}\Xi D_{qp}\tilde{M}\\+\frac{1}{2}\tilde{M}D_{qp}^{\top}\Xi\end{pmatrix}&*\\ A'Q+B'_uY&B'_p\tilde{M}&-Q\end{bmatrix}\prec 0,\tag{17}$$

where $\Xi=\operatorname{diag}\{\xi_i\}\succ 0$. The controller matrix $\bar{K}\in\mathbb{R}^{n_u\times(n_x+n_y)}$ is constructed from $\bar{K}=YQ^{-1}$.

Proof. Given the Lyapunov function

$$V(\tilde{z}_k)=\tilde{z}_k^{\top}P\tilde{z}_k,\ P\succ 0,\tag{18}$$

a sufficient condition for (16) with $\xi_2=\xi>\xi_1=0$ to be g.a.s. is for the inequality

$$\Delta V(\tilde{z}_k)\triangleq V(\tilde{z}_{k+1})-V(\tilde{z}_k)<0,\ \forall k\ge 0,\tag{19}$$

to be satisfied (Hahn, 1958) for all $\tilde{z}_{k,i}$, $\tilde{p}_{k,i}$ that satisfy the sector bounds

$$\tilde{p}_{k,i}\left[\tilde{p}_{k,i}-\xi_i\tilde{q}_{k,i}\right]\le 0,\ \forall k\ge 0,\ i=1,\dots,n_q.\tag{20}$$

The inequality (19) can be written equivalently as the quadratic form

$$\Delta V(\tilde{z}_k)=\begin{bmatrix}\tilde{z}_k\\\tilde{p}_k\end{bmatrix}^{\top}\begin{bmatrix}\tilde{A}^{\top}P\tilde{A}-P&*\\B_p'^{\top}P\tilde{A}&B_p'^{\top}PB'_p\end{bmatrix}\begin{bmatrix}\tilde{z}_k\\\tilde{p}_k\end{bmatrix}<0,\tag{21}$$

and (20) can be written as

$$\begin{bmatrix}\tilde{z}_k\\\tilde{p}_k\end{bmatrix}^{\top}\begin{bmatrix}0&*\\-\frac{\xi_i}{2}E_{i,i}^{\top}\tilde{C}_{q,i}&\begin{pmatrix}E_{i,i}^{\top}E_{i,i}\\-\frac{\xi_i}{2}E_{i,i}^{\top}D_{qp,i}\\-\frac{\xi_i}{2}D_{qp,i}^{\top}E_{i,i}\end{pmatrix}\end{bmatrix}\begin{bmatrix}\tilde{z}_k\\\tilde{p}_k\end{bmatrix}\le 0,\tag{22}$$

where the subscript i in $\tilde{C}_q$ and D_{qp} denotes their ith row, and E_i is a matrix with all its elements equal to zero but $E_{i,ii}=1$. Application of the S-procedure to (21) and (22) gives the sufficient condition

$$\mathrm{F}_1\triangleq\begin{bmatrix}\tilde{A}^{\top}P\tilde{A}-P&*\\B_p'^{\top}P\tilde{A}+\frac{1}{2}\Xi M\tilde{C}_q&\begin{pmatrix}B_p'^{\top}PB'_p-M\\+\frac{1}{2}\Xi MD_{qp}+\frac{1}{2}D_{qp}^{\top}\Xi M\end{pmatrix}\end{bmatrix}\prec 0,\tag{23}$$

where $M = \mathrm{diag}\{m_i\} \succeq 0$ and $\Xi = \mathrm{diag}\{\xi_i\} \succ 0$. To reduce the order of the matrix inequality (23), rewrite as

$$\begin{aligned} &\mathrm{F}_1 = Q - SR^{-1}S^\top \prec 0,\ S \triangleq \begin{bmatrix} \tilde{A}^\top P \\ B_p'^\top P \end{bmatrix},\ R \triangleq -P, \\ &Q \triangleq \begin{bmatrix} -P & * \\ \frac{1}{2}\Xi M\tilde{C}_q & -M + \frac{1}{2}\Xi MD_{qp} + \frac{1}{2}D_{qp}^\top M\Xi \end{bmatrix}. \end{aligned} \tag{24}$$

Then the Schur complement lemma is applied to give

$$\left.\begin{matrix}\mathrm{F}_1 \prec 0 \\ P \succ 0\end{matrix}\right\} \Leftrightarrow \mathrm{F}_2 \triangleq \begin{bmatrix} -P & * & * \\ \frac{1}{2}\Xi M\tilde{C}_q & \begin{pmatrix} -M + \frac{1}{2}\Xi MD_{qp} \\ +\frac{1}{2}D_{qp}^\top M\Xi \end{pmatrix} & * \\ P\tilde{A} & PB_p' & -P \end{bmatrix}. \tag{25}$$

Then a congruence transformation on F_2 with $X = \mathrm{diag}\{P^{-1}, M^{-1}, P^{-1}\}$ results in

$$\begin{aligned} &\mathrm{F}_2 \prec 0 \Leftrightarrow X^\top \mathrm{F}_2 X \prec 0 \Leftrightarrow \\ &\begin{bmatrix} -Q & * & * \\ \frac{1}{2}\Xi\tilde{C}_q Q & \begin{pmatrix} -\tilde{M} + \frac{1}{2}\Xi D_{qp}\tilde{M} \\ +\frac{1}{2}\tilde{M}D_{qp}^\top \Xi \end{pmatrix} & * \\ \tilde{A}Q & B_p'\tilde{M} & -Q \end{bmatrix} \prec 0, \end{aligned} \tag{26}$$

where $Q \triangleq P^{-1} = Q^\top \succ 0$, and $\tilde{M} \triangleq M^{-1} \succ 0$. The implementation of the congruence transformation replaces the requirement $M \succeq 0$ with $M \succ 0$. Substituting (15) into (26) and defining $Y \triangleq \bar{K}Q$ results in the LMI (17). □

Lyapunov stability is established when (17) is nonstrict, i.e., $\mathrm{SF}_{[0,\xi]} \preceq 0$. LMI (17) has $m = \frac{(n_x+n_y)(n_x+n_y+1)}{2} + n_q + n_u(n_x+n_y)$ associated variables and $\mathrm{SF}_{[0,\xi]} \in \mathbb{R}^{n \times n}$, where $n = 2(n_x + n_y) + n_q$. LMI feasibility criteria are computed using semidefinite programming (SDP). The complexity upper bound for decreasing the duality gap of an SDP by a factor of ε^{-1} is known to be $\mathcal{O}((mn + m^2)n^{5/2}\log(\varepsilon))$ (Legat, 2020; Monteiro & Todd, 2000; Monteiro & Zanjácomo, 1999). Assuming a sector bound $[0, \xi]$, $\xi > 0$, results in significantly simpler matrix inequality criteria compared to the general case of a sector bound $[\xi_1, \xi_2]$, $\xi_2 > \xi_1$. LMI criteria for systems of the same structure but with a nonlinear operator sector bounded in $[-1, 1]$ are derived in Section 4.2, which are similar in complexity as for the sector bound $[0, \xi]$, $\xi > 0$, but have different matrix sparsity.

4.2. Full dynamic state error feedback controller for a luré system sector bounded in $[-1, 1]$

This section considers the design of a full dynamic state error feedback controller that stabilizes the closed-loop dynamics. A stabilizing controller is not unique.

Theorem 2. *A sufficient condition for the global asymptotic stability of the controlled system (14) is the existence of $Q \in \mathbb{S}_{++}^{n_x+n_y}$, $\tilde{M} \in \mathbb{D}_{++}^{n_q}$, and $Y \in \mathbb{R}^{n_u \times (n_x+n_y)}$ that satisfy the LMI*

$$\mathrm{SF} \triangleq \begin{bmatrix} -Q & * & * & * \\ 0 & -\tilde{M} & * & * \\ A'Q + B_u'Y & B_p'\tilde{M} & -Q & * \\ C_q'Q + D_{qu}Y & D_{qp}\tilde{M} & 0 & -\tilde{M} \end{bmatrix} \prec 0. \tag{27}$$

The controller matrix $\bar{K} \in \mathbb{R}^{n_u \times (n_x+n_y)}$ is constructed from $\bar{K} = YQ^{-1}$.

Proof. Given the Lyapunov function (18) a sufficient condition for (14) to be g.a.s. is for the inequality (19) to be satisfied (Hahn, 1958) for all $\tilde{z}_{k,i}$, $\tilde{p}_{k,i}$ that satisfy the sector boundedness property

$$\tilde{p}_{k,i}^\top \tilde{p}_{k,i} \le \tilde{q}_{k,i}^\top \tilde{q}_{k,i},\ \forall k \ge 0,\ i = 1, \ldots, n_q. \tag{28}$$

The inequality (19) can be written equivalently as the quadratic form (21) and (28) can be written as

$$\begin{bmatrix} \tilde{z}_k \\ \tilde{p}_k \end{bmatrix}^\top \begin{bmatrix} -\tilde{C}_{q,i}^\top \tilde{C}_{q,i} & * \\ -D_{qp,i}^\top \tilde{C}_{q,i} & E_i - D_{qp,i}^\top D_{qp,i} \end{bmatrix} \begin{bmatrix} \tilde{z}_k \\ \tilde{p}_k \end{bmatrix} \le 0. \tag{29}$$

Application of the S-procedure to (21) and (29) gives the sufficient condition

$$\mathrm{F}_1 \triangleq \begin{bmatrix} \begin{pmatrix} \tilde{A}^\top P\tilde{A} - P \\ +\tilde{C}_q^\top M\tilde{C}_q \end{pmatrix} & * \\ \begin{pmatrix} B_p'^\top P\tilde{A} \\ +D_{qp}^\top M\tilde{C}_q \end{pmatrix} & \begin{pmatrix} B_p'^\top PB_p' - M \\ +D_{qp}^\top MD_{qp} \end{pmatrix} \end{bmatrix} \prec 0, \tag{30}$$

where $M = \mathrm{diag}\{m_i\} \succeq 0$. In (30), higher order terms such as $\tilde{A}^\top P\tilde{A}$ are present. To reduce the order of (30), rewrite as

$$\begin{aligned} &\mathrm{F}_1 = Q - SR^{-1}S^\top \prec 0, \\ &Q = R \triangleq \begin{bmatrix} -P & 0 \\ 0 & -M \end{bmatrix},\quad S \triangleq \begin{bmatrix} \tilde{A}^\top P & \tilde{C}_q^\top M \\ B_p'^\top P & D_{qp}^\top M \end{bmatrix}. \end{aligned} \tag{31}$$

Then the Schur complement lemma is applied to give

$$\left.\begin{matrix}\mathrm{F}_1 \prec 0 \\ M \succ 0\end{matrix}\right\} \Leftrightarrow \mathrm{F}_2 \triangleq \begin{bmatrix} -P & * & * & * \\ 0 & -M & * & * \\ P\tilde{A} & PB_p' & -P & * \\ M\tilde{C}_q & MD_{qp} & 0 & -M \end{bmatrix} \prec 0. \tag{32}$$

In order for the Schur complement lemma to hold, M must be positive definite rather than positive semidefinite. Then a congruence transformation on F_2 with $X = \mathrm{diag}\{P^{-1}, M^{-1}, P^{-1}, M^{-1}\}$ results in

$$\mathrm{F}_2 \prec 0 \Leftrightarrow X^\top \mathrm{F}_2 X \prec 0 \Leftrightarrow \begin{bmatrix} -Q & * & * & * \\ 0 & -\tilde{M} & * & * \\ \tilde{A}Q & B_p'\tilde{M} & -Q & * \\ \tilde{C}_q Q & D_{qp}\tilde{M} & 0 & -\tilde{M} \end{bmatrix} \prec 0, \tag{33}$$

where $Q \triangleq P^{-1} = Q^\top \succ 0$, and $\tilde{M} \triangleq M^{-1} \succ 0$. Substituting (15) into (33) and defining $Y \triangleq \bar{K}Q$ results in the LMI (27). □

The size of (27) is $\mathrm{SF} \in \mathbb{R}^{2(n_x+n_y+n_q)\times 2(n_x+n_y+n_q)}$, which has $\frac{(n_x+n_y)(n_x+n_y+1)}{2} + n_q + n_u(n_x+n_y)$ associated variables. For an output controllable system, a dynamic output feedback controller would be $\bar{K} = \begin{bmatrix} K_p C_y & K_i \end{bmatrix}$. Due to the structural constraint in $\bar{K}$ imposed by C_y, the problem is nonconvex. Ways to solve this control synthesis problem are discussed in a recent paper (Nikolakopoulou et al., 2021).

4.3. Nonlinear dynamic feedback control for a Luré system bounded in $[-1, 1]$

Consider a nonlinear full dynamic state feedback control law $\tilde{u}_k = K_p\tilde{x}_k + K_i\tilde{I}_k + K_n\tilde{p}_k = \bar{K}\tilde{z}_k + K_n\tilde{p}_k$, $\bar{K} = [K_p\ K_i]$, and rewrite (13) as

$$\begin{aligned} \tilde{z}_{k+1} &= \tilde{A}\tilde{z}_k + \tilde{B}_p\tilde{p}_k, \\ \tilde{q}_k &= \tilde{C}_q\tilde{z}_k + \tilde{D}_{qp}\tilde{p}_k, \\ \tilde{p}_{k,i} &= \tilde{\phi}_i(\tilde{q}_{k,i}),\ |\tilde{\phi}_i(\sigma)| \le |\sigma|, \end{aligned} \tag{34}$$

where

$$\tilde{B}_p \triangleq B_p' + B_u'K_n,\ \tilde{D}_{qp} \triangleq D_{qp} + D_{qu}K_n, \tag{35}$$

and all other matrices are defined as in (15).

Theorem 3. *A sufficient condition for the global asymptotic stability of the controlled system (34) is the existence of $Q \in \mathbb{S}_{++}^{n_x+n_y}$, $\tilde{M}$*

$\in \mathbb{D}^{n_q}_{++}$, $Y_1 \in \mathbb{R}^{n_u\times(n_x+n_y)}$, and $Y_2 \in \mathbb{R}^{n_u\times n_q}$ that satisfy the LMI

$$\mathrm{SF}_{\mathrm{nl}} \triangleq \begin{bmatrix} -Q & * & * & * \\ 0 & -\tilde{M} & * & * \\ A'Q + B'_uY_1 & B'_p\tilde{M} + B'_uY_2 & -Q & * \\ C'_qQ + D_{qu}Y_1 & D_{qp}\tilde{M} + D_{qu}Y_2 & 0 & -\tilde{M} \end{bmatrix} \prec 0. \tag{36}$$

The controller matrices $\bar{K} \in \mathbb{R}^{n_u\times(n_x+n_y)}$ and $K_n \in \mathbb{R}^{n_u\times n_q}$ are constructed from $\bar{K} = Y_1Q^{-1}$ and $K_n = Y_2\tilde{M}^{-1}$ respectively.

Proof. Given the Lyapunov function (18), a sufficient condition for (34) to be g.a.s. is for the inequality (19) to be satisfied for all $\tilde{z}_{k,i}$, $\tilde{p}_{k,i}$ that satisfy the sector bounds (28). Inequalities (19) and (28) can be written as their equivalent quadratic forms

$$\Delta V(\tilde{z}_k) = \begin{bmatrix} \tilde{z}_k \\ \tilde{p}_k \end{bmatrix}^\top \begin{bmatrix} \tilde{A}^\top P\tilde{A} - P & * \\ \tilde{B}_p^\top P\tilde{A} & \tilde{B}_p^\top P\tilde{B}_p \end{bmatrix} \begin{bmatrix} \tilde{z}_k \\ \tilde{p}_k \end{bmatrix} < 0 \tag{37}$$

and

$$\begin{bmatrix} \tilde{z}_k \\ \tilde{p}_k \end{bmatrix}^\top \begin{bmatrix} -\tilde{C}_{q,i}^\top\tilde{C}_{q,i} & * \\ -\tilde{D}_{qp,i}^\top\tilde{C}_{q,i} & E_i - \tilde{D}_{qp,i}^\top\tilde{D}_{qp,i} \end{bmatrix} \begin{bmatrix} \tilde{z}_k \\ \tilde{p}_k \end{bmatrix} \le 0. \tag{38}$$

Application of the S-procedure to (37) and (38) gives the sufficient condition for the g.a.s. of (34)

$$\mathrm{F}_1 \triangleq \begin{bmatrix} \begin{pmatrix} \tilde{A}^\top P\tilde{A} - P \\ +\tilde{C}_q^\top M\tilde{C}_q \end{pmatrix} & * \\ \tilde{B}_p^\top P\tilde{A} + \tilde{D}_{qp}^\top M\tilde{C}_q & \begin{pmatrix} \tilde{B}_p^\top P\tilde{B}_p - M \\ +\tilde{D}_{qp}^\top M\tilde{D}_{qp} \end{pmatrix} \end{bmatrix} \prec 0, \tag{39}$$

where $M = \mathrm{diag}\{m_i\} \succeq 0$. To reduce the order of the matrix inequality (39), rewrite as

$$\begin{aligned} &\mathrm{F}_1 = Q - SR^{-1}S^\top \prec 0, \\ &Q = R \triangleq \begin{bmatrix} -P & 0 \\ 0 & -M \end{bmatrix}, \; S \triangleq \begin{bmatrix} \tilde{A}^\top P & \tilde{C}_q^\top M \\ \tilde{B}_p^\top P & \tilde{D}_{qp}^\top M \end{bmatrix}. \end{aligned} \tag{40}$$

Then the Schur complement lemma is applied to give

$$\left.\begin{matrix} \mathrm{F}_1 \prec 0 \\ M \succ 0 \end{matrix}\right\} \Leftrightarrow \mathrm{F}_2 \triangleq \begin{bmatrix} -P & * & * & * \\ 0 & -M & * & * \\ P\tilde{A} & P\tilde{B}_p & -P & * \\ M\tilde{C}_q & M\tilde{D}_{qp} & 0 & -M \end{bmatrix} \prec 0. \tag{41}$$

Then a congruence transformation on F_2 with $X = \mathrm{diag}\{P^{-1}, M^{-1}, P^{-1}, M^{-1}\}$ results in

$$\mathrm{F}_2 \prec 0 \Leftrightarrow X^\top \mathrm{F}_2 X \prec 0 \Leftrightarrow \begin{bmatrix} -Q & * & * & * \\ 0 & -\tilde{M} & * & * \\ \tilde{A}Q & \tilde{B}_p\tilde{M} & -Q & * \\ \tilde{C}_qQ & \tilde{D}_{qp}\tilde{M} & 0 & -\tilde{M} \end{bmatrix} \prec 0, \tag{42}$$

where $Q \triangleq P^{-1} = Q^\top \succ 0$, and $\tilde{M} \triangleq M^{-1} \succ 0$. Substituting (15) and (35) into (42) and defining $Y_1 \triangleq \bar{K}Q$ and $Y_2 \triangleq K_n\tilde{M}$ results in the LMI (36). □

The size of (36) is $\mathrm{SF}_{\mathrm{nl}} \in \mathbb{R}^{2(n_x+n_y+n_q)\times 2(n_x+n_y+n_q)}$, which has $\frac{(n_x+n_y)(n_x+n_y+1)}{2} + n_q + n_u(n_x + n_y + n_q)$ associated variables. In a practical setting, the control law $\tilde{u}_k = K_p\tilde{x}_k + K_i\tilde{I}_k + K_n\tilde{p}_k$ would be implemented as $u_k = K_p(x_k - x_{ss}) + K_iI_k + K_n(p_k - p_{ss})$. While the desired reference value x_{ss} is known, a practical limitation is that implementation requires the value of p_{ss} which is not known due to the unknown nature of I_{ss}.

5. State observer design

Consider a model of the similar structure as (12), with the inclusion of the output equation:

$$\begin{aligned} x_{k+1} &= Ax_k + B_pp_k + B_uu_k + B_dd_k + b_c, \quad x(0) = x_0, \\ q_k &= C_qx_k + D_{qp}p_k + D_{qu}u_k + D_{qd}d_k + d_{qc}, \\ y_k &= C_yx_k + D_{yp}p_k + D_{yu}u_k, \\ p_{k,i} &= \phi_i(q_{k,i}), \; |\phi_i(\sigma)| \le |\sigma|, \end{aligned} \tag{43}$$

where y_k is the vector of outputs with available measurements. The corresponding observer dynamics are

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + B_p\hat{p}_k + B_uu_k + B_dd_k + b_c + L_1(\hat{y}_k - y_k), \\ \hat{q}_k &= C_q\hat{x}_k + D_{qp}\hat{p}_k + D_{qu}u_k + D_{qd}d_k + d_{qc} \\ &\quad + L_2(\hat{y}_k - y_k), \\ \hat{y}_k &= C_y\hat{x}_k + D_{yp}\hat{p}_k + D_{yu}u_k, \\ \hat{p}_{k,i} &= \phi_i(\hat{q}_{k,i}), \; |\phi_i(\sigma)| \le |\sigma|, \end{aligned} \tag{44}$$

initialized as $\hat{x}(0) = \hat{x}_0$. The observer error dynamics are

$$\begin{aligned} \hat{x}_{k+1} - x_{k+1} &= (A + L_1C_y)(\hat{x}_k - x_k) + (B_p + L_1D_{yp})(\hat{p}_k - p_k), \\ \hat{q}_k - q_k &= (C_q + L_2C_y)(\hat{x}_k - x_k) \\ &\quad + (D_{qp} + L_2D_{yp})(\hat{p}_k - p_k), \\ \hat{y}_k - y_k &= C_y(\hat{x}_k - x_k) + D_{yp}(\hat{p}_k - p_k), \\ \hat{p}_{k,i} - p_{k,i} &= \delta_i(k)(\hat{q}_{k,i} - q_{k,i}), \; |\delta_i(k)| \le 1. \end{aligned} \tag{45}$$

Theorem 4. *A sufficient condition for the observer error dynamics (45) to have a decay rate of at least α is the existence of $P \in \mathbb{S}^{n_x}_{++}$, $M \in \mathbb{D}^{n_q}_{++}$, $W_1 \in \mathbb{R}^{n_x\times n_y}$, and $W_2 \in \mathbb{R}^{n_q\times n_y}$ that satisfy the LMI*

$$\mathrm{SE} \triangleq \begin{bmatrix} \mathrm{SE}_{11} & * & * & * \\ \mathrm{SE}_{21} & \mathrm{SE}_{22} & * & * \\ \mathrm{SE}_{31} & \mathrm{SE}_{32} & \mathrm{SE}_{33} & * \\ \mathrm{SE}_{41} & \mathrm{SE}_{42} & \mathrm{SE}_{43} & \mathrm{SE}_{44} \end{bmatrix} \preceq 0, \tag{46}$$

where

$$\begin{aligned} &\mathrm{SE}_{11} = -(1-\alpha)P, && \mathrm{SE}_{21} = \mathrm{SE}_{43} = 0, \\ &\mathrm{SE}_{22} = \mathrm{SE}_{44} = -M, && \mathrm{SE}_{31} = PA + W_1C_y, \\ &\mathrm{SE}_{32} = PB_p + W_1D_{yp}, && \mathrm{SE}_{33} = -P, \\ &\mathrm{SE}_{41} = MC_q + W_2C_y, && \mathrm{SE}_{42} = MD_{qp} + W_2D_{yp}. \end{aligned} \tag{47}$$

The observer matrices $L_1 \in \mathbb{R}^{n_x\times n_y}$ and $L_2 \in \mathbb{R}^{n_q\times n_y}$ are constructed from $L_1 = P^{-1}W_1$ and $L_2 = M^{-1}W_2$.

Proof. Given a Lyapunov function

$$V(\hat{x}_k - x_k) = (\hat{x}_k - x_k)^\top P(\hat{x}_k - x_k), \; P \succ 0, \tag{48}$$

define

$$\Delta V(\hat{x}_k - x_k) \triangleq V(\hat{x}_{k+1} - x_{k+1}) - V(\hat{x}_k - x_k). \tag{49}$$

If the observer error dynamics satisfy

$$\Delta V(\hat{x}_k - x_k) \le -\alpha V(\hat{x}_k - x_k), \; \forall k \ge 0, \tag{50}$$

for all $\hat{x}_{k,i} - x_{k,i}$, $\hat{p}_{k,i} - p_{k,i}$ that satisfy the sector bounds

$$\begin{aligned} &(\hat{p}_{k,i} - p_{k,i})^\top(\hat{p}_{k,i} - p_{k,i}) \le (\hat{q}_{k,i} - q_{k,i})^\top(\hat{q}_{k,i} - q_{k,i}), \\ &\forall k \ge 0, \; i = 1, \ldots, n_q, \end{aligned} \tag{51}$$

then the decay rate is at least α (Boyd et al., 1994). The inequality (50) can be written equivalently as the quadratic form

$$\hat{\zeta}_k^\top \begin{bmatrix} \hat{A}^\top P\hat{A} - (1-\alpha)P & * \\ \hat{B}_p^\top P\hat{A} & \hat{B}_p^\top P\hat{B}_p \end{bmatrix} \hat{\zeta}_k \le 0, \tag{52}$$

where

$$\hat{\zeta}_k \triangleq \begin{bmatrix} \hat{x}_k - x_k \\ \hat{p}_k - p_k \end{bmatrix}, \; \hat{A} \triangleq A + L_1C_y, \; \hat{B}_p \triangleq B_p + L_1D_{yp}, \tag{53}$$

and (51) can be written equivalently as

$$\hat{\zeta}_k^\top \begin{bmatrix} -\hat{C}_{q,i}^\top \hat{C}_{q,i} & * \\ -\hat{D}_{qp,i}^\top \hat{C}_{q,i} & E_i - \hat{D}_{qp,i}^\top \hat{D}_{qp,i} \end{bmatrix} \hat{\zeta}_k \leq 0, \tag{54}$$

where

$$\hat{C}_q \triangleq C_q + L_2 C_y, \; \hat{D}_{qp} \triangleq D_{qp} + L_2 D_{yp}. \tag{55}$$

Application of the S-procedure to (52) and (54) gives the sufficient condition

$$\mathrm{SE}_1 \triangleq \begin{bmatrix} \begin{pmatrix} \hat{A}^\top P \hat{A} - (1-\alpha)P \\ +\hat{C}_q^\top M \hat{C}_q \end{pmatrix} & * \\ \hat{B}_p^\top P \hat{A} + \hat{D}_{qp}^\top M \hat{C}_q & \begin{pmatrix} \hat{B}_p^\top P \hat{B}_p - M \\ +\hat{D}_{qp}^\top M \hat{D}_{qp} \end{pmatrix} \end{bmatrix} \preceq 0, \tag{56}$$

where $M = \operatorname{diag}\{m_i\} \succeq 0$. In (56), higher order terms such as $\hat{A}^\top P \hat{A}$ are present. To reduce the order of (56) to an LMI, rewrite the matrix inequality as

$$\begin{aligned} &\mathrm{SE}_1 = Q - SR^{-1}S^\top \preceq 0, \; R \triangleq \begin{bmatrix} -P & 0 \\ 0 & -M \end{bmatrix} \\ &Q \triangleq \begin{bmatrix} -(1-\alpha)P & 0 \\ 0 & -M \end{bmatrix}, \; S \triangleq \begin{bmatrix} \hat{A}^\top P & \hat{C}_q^\top M \\ \hat{B}_p^\top P & \hat{D}_{qp}^\top M \end{bmatrix}. \end{aligned} \tag{57}$$

Then

$$\left.\begin{matrix} \mathrm{SE}_1 \preceq 0 \\ M \succ 0 \end{matrix}\right\} \Leftrightarrow \mathrm{SE}_2 \triangleq \begin{bmatrix} -(1-\alpha)P & * & * & * \\ 0 & -M & * & * \\ P\hat{A} & P\hat{B}_p & -P & * \\ M\hat{C}_q & M\hat{D}_{qp} & 0 & -M \end{bmatrix} \preceq 0. \tag{58}$$

Substituting (53) and (55) into (58) and defining $W_1 \triangleq PL_1$ and $W_2 \triangleq ML_2$ results in the LMI (46). □

The size of (46) is $\mathrm{SE} \in \mathbb{R}^{2(n_x+n_q)\times 2(n_x+n_q)}$, which has $\frac{n_x(n_x+1)}{2} + n_q + (n_x+n_q)n_y$ associated variables. The observer that maximizes the decay rate of (45) is given by the GEVP:

$$\begin{aligned} &\max_{P,W,M} \; \alpha \\ &\text{s.t.} \quad M = \operatorname{diag}\{m_i\} \succ 0, \; P = P^\top \succ 0, \; \mathrm{SE} \preceq 0. \end{aligned} \tag{59}$$

6. Output observer design

Consider a model of similar structure as (12), with an additional unmeasured disturbance w_k acting on the system

$$\begin{aligned} x_{k+1} &= Ax_k + B_p p_k + B_u u_k + B_d d_k + B_w w_k + b_c, \\ q_k &= C_q x_k + D_{qp} p_k + D_{qu} u_k + D_{qd} d_k + D_{qw} w_k + d_{qc}, \\ y_k &= C_y x_k + D_{yp} p_k + D_{yu} u_k + D_{yw} w_k, \\ z_k &= C_z x_k + D_{zp} p_k + D_{zw} w_k, \\ p_{k,i} &= \phi_i(q_{k,i}), \quad |\phi_i(\sigma)| \leq |\sigma|, \end{aligned} \tag{60}$$

with initial conditions $x(0) = x_0$, where y_k is the vector of outputs and z_k is the vector to be estimated.[2] The corresponding observer dynamics are

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + B_p \hat{p}_k + B_u u_k + B_d d_k + b_c + L_1(\hat{y}_k - y_k), \\ \hat{q}_k &= C_q \hat{x}_k + D_{qp} \hat{p}_k + D_{qu} u_k + D_{qd} d_k + d_{qc} \\ &\quad + L_2(\hat{y}_k - y_k), \\ \hat{y}_k &= C_y \hat{x}_k + D_{yp} \hat{p}_k + D_{yu} u_k, \\ \hat{z}_k &= C_z \hat{x}_k + D_{zp} \hat{p}_k, \\ \hat{p}_{k,i} &= \phi_i(\hat{q}_{k,i}), \quad |\phi_i(\sigma)| \leq |\sigma|, \end{aligned} \tag{61}$$

initialized as $\hat{x}(0) = \hat{x}_0$. The observer error dynamics are

$$\begin{aligned} \hat{x}_{k+1} - x_{k+1} &= (A+L_1C_y)(\hat{x}_k - x_k) + (B_p + L_1 D_{yp})(\hat{p}_k - p_k) \\ &\quad - (B_w + L_1 D_{yw}) w_k, \\ \hat{q}_k - q_k &= (C_q + L_2 C_y)(\hat{x}_k - x_k) - (D_{qw} + L_2 D_{yw}) w_k \\ &\quad + (D_{qp} + L_2 D_{yp})(\hat{p}_k - p_k), \\ \hat{y}_k - y_k &= C_y(\hat{x}_k - x_k) + D_{yp}(\hat{p}_k - p_k) - D_{yw} w_k, \\ \hat{z}_k - z_k &= C_z(\hat{x}_k - x_k) + D_{zp}(\hat{p}_k - p_k) - D_{zw} w_k, \\ \hat{p}_{k,i} - p_{k,i} &= \delta_i(k)(\hat{q}_{k,i} - q_{k,i}), \; |\delta_i(k)| \leq 1. \end{aligned} \tag{62}$$

The L_2-norm of w_k in discrete time is defined as $\|w\|_2^2 \triangleq \sum_{k=0}^{\infty} w_k^\top w_k$, and the L_2-gain as $\sup_{\|w\|_2 \neq 0} \frac{\|\hat{z}-z\|_2}{\|w\|_2}$.

Theorem 5. *A sufficient condition for the observer error dynamics (62) to have an L_2-gain less than γ is the existence of $P \in \mathbb{S}_{++}^{n_x}$, $M \in \mathbb{D}_{++}^{n_q}$, $W_1 \in \mathbb{R}^{n_x \times n_y}$, and $W_2 \in \mathbb{R}^{n_q \times n_y}$ that satisfy the LMI*

$$\mathrm{OE} \triangleq \begin{bmatrix} \mathrm{OE}_{11} & * & * & * & * \\ \mathrm{OE}_{21} & \mathrm{OE}_{22} & * & * & * \\ \mathrm{OE}_{31} & \mathrm{OE}_{32} & \mathrm{OE}_{33} & * & * \\ \mathrm{OE}_{41} & \mathrm{OE}_{42} & \mathrm{OE}_{43} & \mathrm{OE}_{44} & * \\ \mathrm{OE}_{51} & \mathrm{OE}_{52} & \mathrm{OE}_{53} & \mathrm{OE}_{54} & \mathrm{OE}_{55} \end{bmatrix} \preceq 0, \tag{63}$$

where

$$\begin{aligned} &\mathrm{OE}_{11} = -P + C_z^\top C_z, && \mathrm{OE}_{21} = D_{zp}^\top C_z, \\ &\mathrm{OE}_{22} = -M + D_{zp}^\top D_{zp}, && \mathrm{OE}_{31} = -D_{zw}^\top C_z, \\ &\mathrm{OE}_{32} = -D_{zw}^\top D_{zp}, && \mathrm{OE}_{33} = -\gamma^2 I + D_{zw}^\top D_{zw}, \\ &\mathrm{OE}_{41} = PA + W_1 C_y, && \mathrm{OE}_{42} = PB_p + W_1 D_{yp}, \\ &\mathrm{OE}_{43} = -PB_w - W_1 D_{yw}, && \mathrm{OE}_{44} = -P, \\ &\mathrm{OE}_{51} = MC_q + W_2 C_y, && \mathrm{OE}_{52} = MD_{qp} + W_2 D_{yp}, \\ &\mathrm{OE}_{53} = -MD_{qw} - W_2 D_{yw}, && \mathrm{OE}_{54} = 0, \\ &\mathrm{OE}_{55} = -M. \end{aligned} \tag{64}$$

The observer matrices $L_1 \in \mathbb{R}^{n_x \times n_y}$ and $L_2 \in \mathbb{R}^{n_q \times n_y}$ are constructed from $L_1 = P^{-1} W_1$ and $L_2 = M^{-1} W_2$.

Proof. Given a Lyapunov function (48), define (49). If the observer error dynamics (62) satisfy

$$\Delta V(\hat{x}_k - x_k) + (\hat{z}_k - z_k)^\top (\hat{z}_k - z_k) - \gamma^2 w_k^\top w_k \leq 0, \; \forall k \geq 0, \tag{65}$$

for all $\hat{x}_{k,i} - x_{k,i}$, $\hat{p}_{k,i} - p_{k,i}$ that satisfy the sector bounds (51), then the L_2-gain is less than γ (Boyd et al., 1994). The inequality (65) can be written equivalently as

$$\hat{\zeta}_k^\top \begin{bmatrix} \begin{pmatrix} \hat{A}^\top P \hat{A} \\ -P + C_z^\top C_z \end{pmatrix} & * & * \\ \begin{pmatrix} \hat{B}_p^\top P \hat{A} \\ +D_{zp}^\top C_z \end{pmatrix} & \begin{pmatrix} \hat{B}_p^\top P \hat{B}_p \\ +D_{zp}^\top D_{zp} \end{pmatrix} & * \\ \begin{pmatrix} -\hat{B}_w^\top P \hat{A} \\ -D_{zw}^\top C_z \end{pmatrix} & \begin{pmatrix} -\hat{B}_w^\top P \hat{B}_p \\ -D_{zw} D_{zp} \end{pmatrix} & \begin{pmatrix} \hat{B}_w^\top P \hat{B}_w - \gamma^2 I \\ +D_{zw}^\top D_{zw} \end{pmatrix} \end{bmatrix} \hat{\zeta}_k \leq 0, \tag{66}$$

where

$$\begin{aligned} &\hat{\zeta}_k \triangleq [\hat{x}_k - x_k \;\; \hat{p}_k - p_k \;\; w_k]^\top, \; \hat{A} \triangleq A + L_1 C_y, \\ &\hat{B}_p \triangleq B_p + L_1 D_{yp}, \; \hat{B}_w \triangleq B_w + L_1 D_{yw}, \end{aligned} \tag{67}$$

and (51) can be written equivalently as

$$\hat{\zeta}_k^\top \begin{bmatrix} -\hat{C}_{q,i}^\top \hat{C}_{q,i} & * & * \\ -\hat{D}_{qp,i}^\top \hat{C}_{q,i} & -\hat{D}_{qp,i}^\top \hat{D}_{qp,i} + E_i & * \\ \hat{D}_{qw,i}^\top \hat{C}_{q,i} & \hat{D}_{qw,i}^\top \hat{D}_{qp,i} & -\hat{D}_{qw,i}^\top \hat{D}_{qw,i} \end{bmatrix} \hat{\zeta}_k \leq 0, \tag{68}$$

[2] State estimation is the special case where $C_z = I$, $D_{zp} = 0$, and $D_{zw} = 0$.

where

$$\begin{aligned}&\hat{C}_q \triangleq C_q + L_2 C_y,\ \hat{D}_{qp} \triangleq D_{qp} + L_2 D_{yp},\\ &\hat{D}_{qw} \triangleq D_{qw} + L_2 D_{yw}.\end{aligned} \tag{69}$$

Application of the S-procedure to (66) and (68) gives the sufficient condition

$$\mathrm{OE}_1 \triangleq \begin{bmatrix} \mathrm{OE}_{1,11} & * & * \\ \mathrm{OE}_{1,21} & \mathrm{OE}_{1,22} & * \\ \mathrm{OE}_{1,31} & \mathrm{OE}_{1,32} & \mathrm{OE}_{1,33} \end{bmatrix} \preceq 0, \tag{70}$$

where

$$\begin{aligned}
\mathrm{OE}_{1,11} &= \hat{A}^\top P \hat{A} - P + C_z^\top C_z + \hat{C}_q^\top M \hat{C}_q,\\
\mathrm{OE}_{1,21} &= \hat{B}_p^\top P \hat{A} + D_{zp}^\top C_z + \hat{D}_{qp}^\top M \hat{C}_q,\\
\mathrm{OE}_{1,22} &= \hat{B}_p^\top P \hat{B}_p + D_{zp}^\top D_{zp} + \hat{D}_{qp}^\top M \hat{D}_{qp} - M,\\
\mathrm{OE}_{1,31} &= -\hat{B}_w^\top P \hat{A} - D_{zw}^\top C_z - \hat{D}_{qw}^\top M \hat{C}_q,\\
\mathrm{OE}_{1,32} &= -\hat{B}_w^\top P \hat{B}_p - D_{zw}^\top D_{zp} - \hat{D}_{qw}^\top M \hat{D}_{qp},\\
\mathrm{OE}_{1,33} &= \hat{B}_w^\top P \hat{B}_w + D_{zw}^\top D_{zw} + \hat{D}_{qw}^\top M \hat{D}_{qw} - \gamma^2 I,
\end{aligned} \tag{71}$$

and $M = \mathrm{diag}\{m_i\} \succeq 0$. To reduce the order of (70) to an LMI, rewrite the matrix inequality as

$$\begin{aligned}
&\mathrm{OE}_1 = Q - SR^{-1}S^\top \preceq 0,\\
&Q \triangleq \begin{bmatrix} -P + C_z^\top C_z & * & * \\ D_{zp}^\top C_z & D_{zp}^\top D_{zp} - M & * \\ -D_{zw}^\top C_z & -D_{zw}^\top D_{zp} & D_{zw}^\top D_{zw} - \gamma^2 I \end{bmatrix},\\
&S \triangleq \begin{bmatrix} \hat{A}^\top P & \hat{C}_q^\top M \\ \hat{B}_p^\top P & \hat{D}_{qp}^\top M \\ -\hat{B}_w^\top P & -\hat{D}_{qw}^\top M \end{bmatrix},\quad R \triangleq \begin{bmatrix} -P & 0 \\ 0 & -M \end{bmatrix}.
\end{aligned} \tag{72}$$

Then

$$\left.\begin{matrix}\mathrm{OE}_1 \preceq 0 \\ M \succ 0\end{matrix}\right\} \Leftrightarrow \mathrm{OE}_2 \triangleq \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \preceq 0. \tag{73}$$

Substituting (67) and (69) into (73) and defining $W_1 \triangleq PL_1$ and $W_2 \triangleq ML_2$ results in the LMI (63) and (64). □

The size of (63) is $\mathrm{OE} \in \mathbb{R}^{(2n_x+2n_q+n_w)\times(2n_x+2n_q+n_w)}$, which has $\frac{n_x(n_x+1)}{2} + n_q + (n_x + n_q)n_y$ associated variables. The observer that minimizes the upper bound of the L_2-gain of (62) is given by the solution of the eigenvalue problem (EVP):

$$\begin{aligned}&\min_{P,W,M} \gamma^2\\ &\text{s.t.}\quad M = \mathrm{diag}\{m_i\} \succ 0,\ P = P^\top \succ 0,\ \mathrm{OE} \preceq 0.\end{aligned} \tag{74}$$

7. Case studies

7.1. Numerical example

Here the methodology is implemented for a numerical example with one input (u), two states (x_1, x_2), and one output (x_2). The system has the structure of a two-layer feedforward neural network with three nodes in the first layer and two nodes in the second layer.

Table 2
Controller and observer matrices for the numerical example.

Description	Value
Controller	$\bar{K} = [0.0386\quad 0.0283\quad 0.0340]$
Observer 1	$L_1 = [0.2544\quad 0.3963]^\top$
Observer 2	$L_2 = \begin{bmatrix} 0.1187 \\ -0.0360 \\ -0.2676 \\ -0.0611 \\ -0.0077 \end{bmatrix}$
PI controller	$K_c = -0.0017$, $\tau_I = 0.03$

7.1.1. Problem formulation

The model approximating the system is described by (6) with

$$\begin{aligned}
&A = \begin{bmatrix} -0.5035 & 0.0550 \\ -0.0094 & -0.3834 \end{bmatrix},\quad B_u = \begin{bmatrix} 11 \\ -15 \end{bmatrix},\\
&B_p = \begin{bmatrix} \mathbf{0}_{1\times3} & 0.2244 & -0.0800 \\ \mathbf{0}_{1\times3} & 0.3573 & 0.7400 \end{bmatrix},\quad b_c = \mathbf{0}_{2\times1},\\
&C_q = \begin{bmatrix} -0.0756 & -0.0746 \\ -0.0675 & 0.0766 \\ 0.1200 & 0.2000 \\ \mathbf{0}_{2\times1} & \mathbf{0}_{2\times1} \end{bmatrix},\quad d_{qc} = \begin{bmatrix} 0.7408 \\ -0.2018 \\ -1.800 \\ \mathbf{0}_{2\times1} \end{bmatrix},\\
&D_{qp} = \begin{bmatrix} \mathbf{0}_{3\times1} & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times2} \\ 0.5000 & 0.2500 & 0.6400 & \mathbf{0}_{1\times2} \\ 0.3340 & 0.9000 & 0.0734 & \mathbf{0}_{1\times2} \end{bmatrix},\quad D_{qu} = \begin{bmatrix} -1.000 \\ 0.3560 \\ 3.000 \\ \mathbf{0}_{2\times1} \end{bmatrix},
\end{aligned} \tag{75}$$

where ϕ is the tanh function. The true process has the structure (6) with the inclusion of an unmeasured disturbance w such that

$$\begin{aligned}
x_{k+1} &= Ax_k + B_p p_k + B_u u_k + B_w w_k + b_c,\\
q_k &= C_q x_k + D_{qp} p_k + D_{qu} u_k + D_{qw} w_k + d_{qc},\\
p_{k,i} &= \phi_i(q_{k,i}),\ i = 1,\ldots,5,
\end{aligned} \tag{76}$$

where $B_w = [-3.0\ 0.1]^\top$, $D_{qw} = \mathbf{0}_{5\times1}$, and the rest of the matrices are defined in (75).

Only the second state, x_2, is measured and controlled to a setpoint. The estimated value for the first state, $\hat{x}_1$, is used in the LMI-based feedback controller implementation such that

$$\begin{aligned}
&u_k = \bar{K}[\bar{x}_k\quad I_k]^\top,\quad \bar{x}_k \triangleq [\hat{x}_{1,k}\quad x_{2,k} - r_k]^\top,\\
&I_{k+1} = I_k + x_{2,k} - r_k,\quad I(0) = 0.
\end{aligned} \tag{77}$$

A proportional–integral (PI) controller for the pair u–x_2 was implemented for comparison purposes.

7.1.2. Results and discussion

The stabilizing controller was obtained by solving the LMI (27), while the observer that maximizes the decay rate of the state observer error dynamics was obtained from (59). The resulting maximum decay rate is $\alpha = 0.7393$. The results are given in Table 2 and Fig. 1. The PI controller was designed using Internal Model Control (IMC) (Morari & Zafiriou, 1989) with the controller parameters given in Table 2.

The unmeasured disturbance acting on the process (Fig. 1a) results in inaccurate estimates for the first state x_1, as expected (Fig. 1c). The inaccurate estimate does not result in an offset in the control of the second state x_2 (Fig. 1d), since both controllers have integral action. The second state is tightly controlled to the setpoint by both the LMI-based and PI controllers (Fig. 1d). The PI controller was designed to have nearly as fast of a closed-loop response speed in the controlled state x_2 as the LMI-based controller (Fig. 1d), which results in some erratic behavior in both states (Figs. 1cd). Designing the PI control to be less aggressive

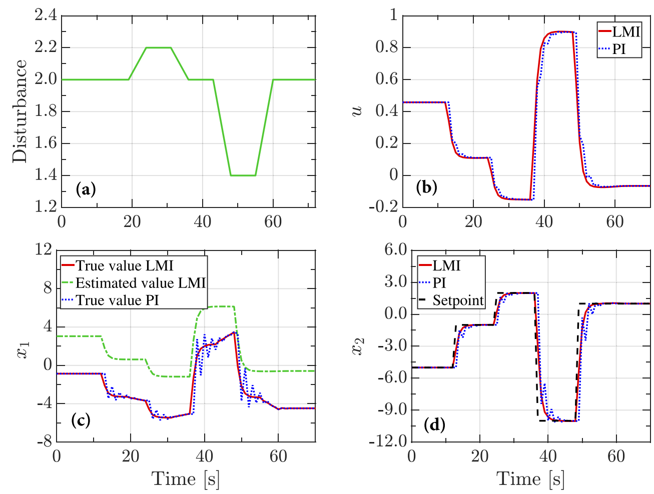

Fig. 1. Closed-loop system behavior during step changes in the setpoint and under the presence of unmeasured ramp disturbance. (a) Unmeasured disturbance. (b) Control input. (c) State 1. (d) State 2.

would reduce the erratic behavior in the states x_1 and x_2 but would result in the closed-loop response in the controlled state x_2 to be more sluggish than the LMI-based controller.

7.2. pH control

This section demonstrates the theoretical results for a pH-neutralization process, which is both highly nonlinear and industrially important (Gustafsson, Skrifvars, Sandström, & Waller, 1995). The system is described by a two-layer neural network with four nodes in the first layer and two nodes in the second layer. The system has two inputs (q_1, q_2), two states (c_1, c_2), and one output (c_2).

7.2.1. Problem formulation

pH neutralization is often performed in multiple mixing tanks in series to gradually increase or decrease the pH to the desired levels (Riggs & Rhinehart, 1990). The pH neutralization reaction is $H^+ + OH^- \rightleftharpoons H_2O$ and the incoming stream can be either a base or an acid.

Here, our goal is to control the pH of an incoming acid stream of pH 3 using two tanks in series (Fig. 2a). We control the pH of that stream by using a base of pH 11. Mass balances of species H^+ and OH^- in the tanks describe the time evolution of the species. Alternatively, the molar concentration of the excess acid in the ith tank of $c_i = c_{H^+} - c_{OH^-}$ is a conserved quantity without reaction terms (Faanes & Skogestad, 2004; Gustafsson et al., 1995). The first-principles equations that describe the process of two tanks in series are

$$\begin{aligned}\tfrac{dc_1}{dt} &= \tfrac{1}{V_1}(c_{in}q_{in} + c_{b,1}q_{b,1} - c_1q_1),\ q_1 = q_{in} + q_{b,1},\\ \tfrac{dc_2}{dt} &= \tfrac{1}{V_2}(c_1q_1 + c_{b,2}q_{b,2} - c_2q_2),\ q_2 = q_1 + q_{b,2},\end{aligned} \tag{78}$$

where V_i is the volume, c_{in} is the excess acid in the inlet stream, q_{in} is the flowrate of the incoming acid stream, $q_{b,i}$ is the flowrate of the base, $c_{b,i}$ is the excess acid in the base stream, q_i is the outflow from the tank, and the subscript i refers to the quantity corresponding to the ith tank. The flowrate of the incoming acid stream q_{in} is modeled as an unmeasured disturbance. The remaining variables are process parameters given in Table 3. The pH is related to the excess acid concentration by $\mathrm{pH}_i = -\log_{10}(c_i/2 + \sqrt{10^{-14} + c_i^2/4})$.

Table 3
PH-neutralization process parameters.

Parameter	V_1, V_2	c_{in}	q_{in}	$c_{b,1}$, $c_{b,2}$
Value	1	10^{-3}	5	-10^{-3}
Units	L	mol/L	L/s	mol/L

System identification was carried out by randomly perturbing the system inputs within their expected operating range and recording measurements every 1 s. The perturbations were uniform random sequences. Data obtained from simulations were used to train the two-layer neural network,

$$\begin{aligned}x_{k+1} &= Ax_k + B_u u_k + B_w w_k + B_p p_k + b_c,\\ q_k &= C_q x_k + D_{qu} u_k + D_{qw} w_k + D_{qp} p_k + d_{qc},\\ p_{k,i} &= \tanh(q_{k,i}),\ i = 1, \ldots, 6,\end{aligned} \tag{79}$$

where

$$x_k = [c_{1,k}\ \ c_{2,k}]^\top,\ u_k = [q_{1,k}\ \ q_{2,k}]^\top,\ w_k = q_{in,k}.$$

To account for model-plant mismatch as would occur in a real experimental implementation, the physical plant is represented by the first-principles equations (78) under the presence of the unmeasured disturbance q_{in}. The pH is measured only at the outlet of the second tank, and Gaussian random noise was assumed to act on the measurement of $c_{2,m} = c_2 + \mathcal{N}(0, \sigma_{c_2})$, where the standard deviation σ_{c_2} is computed through $\sigma_{c_2} = \frac{dc_2}{dpH_2}\sigma_{pH_2}$. We consider $3\sigma_{pH_2} = 0.05$, which is a realistic value for pH probe measurements (Anon, 2010).

The estimate of the first state $\hat{c}_{1,k}$ and the measurement of the second state $c_{2,m,k}$ were used in the feedback controller such that

$$\begin{aligned}u_k &= \bar{K}[\bar{x}_k\ \ I_k]^\top,\ \bar{x}_k \triangleq [\hat{c}_{1,k}\ \ c_{2,m,k}]^\top - r_k,\\ I_{k+1} &= I_k + \bar{x}_k,\ I(0) = 0,\end{aligned} \tag{80}$$

where r_k is the reference at sampling instance k. The estimate of the first state $\hat{c}_{1,k}$ is used in (80) since there is no available pH measurement in the first tank. The estimate will be inaccurate when an unmeasured disturbance is affecting the system. Then, the first tank will not be exactly controlled at setpoint. However, exact setpoint tracking is not required in the first tank, as its role is to dampen in the effects of disturbances so that exact setpoint tracking is achieved in the second tank. In (80), the measurement for the second state is used due to its higher accuracy compared to its corresponding observed value when model-plant mismatch is introduced. An observer-based PI controller was also implemented for comparison purposes. The observer was introduced in the PI controller to inform the value of the pH in the first tank. Lastly, the PI controller error was calculated from the pH measurement (or observed value) and not the excess acid.

7.2.2. Results and discussion

The stabilizing controller was obtained by solving the nonstrict version of LMI (27), SF $\preceq$ 0, implying Lyapunov stability. The observer that minimizes the upper bound for the L_2-gain of the output c_2 to the unmeasured disturbance q_{in} was obtained from (74), resulting in $\gamma^2_{min} = 0.0034$. The results are given in Table 4 and Fig. 2. A stabilizing controller is not unique, which provides some extra degrees of freedom that can be used by the designer. The PI controller was tuned using the IMC method (Morari & Zafiriou, 1989) and the controller parameters are given in Table 4.

In the physical system, up to a 5% hourly pump drift can occur in the incoming acid stream flowrate disturbance. This upper bound is adjusted to 5-minute intervals as shown in Fig. 2b. Simultaneously to the unmeasured disturbance, the setpoints in the first and second tank were varied consecutively. The excess

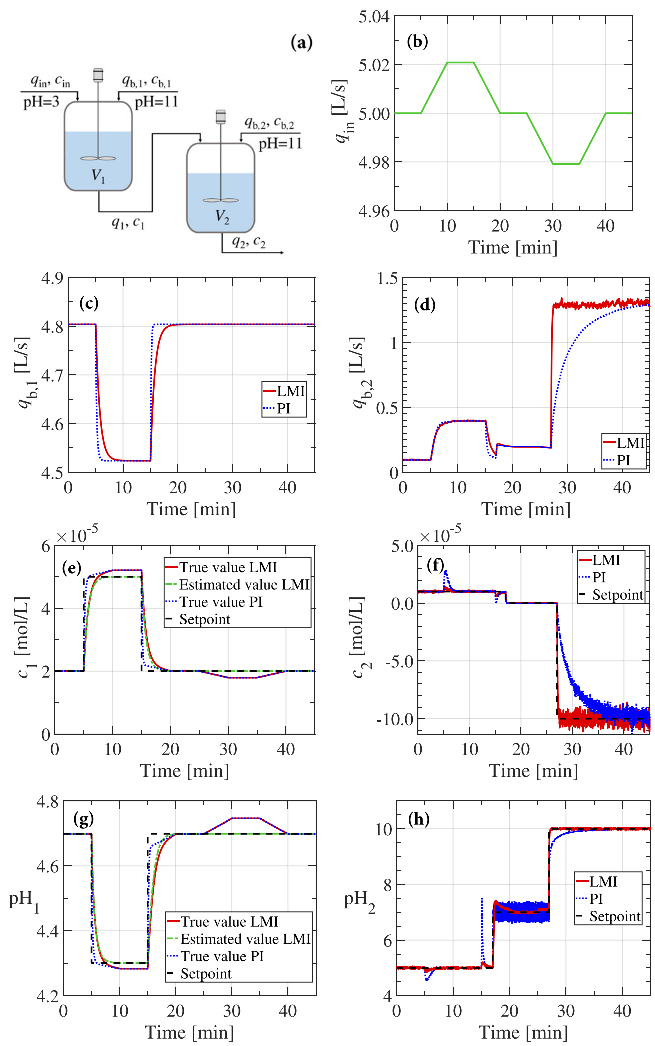

Fig. 2. Closed-loop system behavior during step changes in the setpoint and under the presence of unmeasured ramp disturbance in the incoming acid stream flowrate q_{in}. (a) pH-neutralization process flow diagram. (b) Unmeasured disturbance in the incoming acid stream flowrate q_{in}. (c) Incoming base flowrate in tank 1. (d) Incoming base flowrate in tank 2. (e) Excess acid in tank 1. (f) Excess acid measurement in tank 2. (g) pH in tank 1. (h) pH measurement in tank 2.

Table 4
Controller and observer matrices for the pH control case study.

Description	Value
Controller	$\bar{K} = \begin{bmatrix} 0.1685 & 0 & 200 & 0 \\ 0.1 & 0.3426 & 10 & 1500 \end{bmatrix}$
Observer 1	$L_1 = [0 \quad -8.9\times 10^{-5}]^\top$
Observer 2	$L_2 = [-5 \quad -1 \quad -2100 \quad 1400 \quad 3 \quad 1100]^\top \times 10^{-4}$
PI controller	$K_c = [2.42\times 10^{-4} \quad 4\times 10^{-5}]^\top$, $\tau_I = [0.0033 \quad 0.0040]^\top$

acid c_1 tracks the setpoint (Fig. 2e) with some offset due to a small difference between the estimated and the true value of c_1 for both the LMI-based and PI controllers. As already mentioned, this offset is expected due to the unmeasured disturbance. Because the purpose of the first tank is to act as a disturbance dampener, zero offset is not a performance requirement for that tank. The LMI-based controller is slightly slower at tracking the pH in the first tank (Figs. 2eg) compared to the PI controller. Deviations from the setpoint in c_2 are observed around the times when the setpoint in the first tank changes (Fig. 2f). The deviations are bigger for the PI controller. The excess acid in the two tanks operates at different orders of magnitude, and increasing the excess acid in the first tank has a large effect on the excess acid in the second tank. Such a setpoint change would not occur in the real system but is employed here to illustrate the high nonlinearity of the system. As seen in Figs. 2df, the LMI-based controller effectively rejects disturbances and rapidly responds to changes in tank 2's c_2 setpoint within the constraint bounds of ± 1 pH unit. However, the PI controller is slower at responding to setpoint changes around values of pH quite smaller and larger than 7 in the second tank. The PI controller results in large variations in the pH value around pH 7. These variations could be resolved with different tuning around that setpoint which would then result in worse performance in other regions of the operation. This behavior attests to the process's strong nonlinearity.

8. Conclusions

This article describes design criteria for dynamic state feedback controllers and state and output observers with guaranteed properties for systems described by DANNs. Starting from an available DANN model of a dynamical process, the DANN is reformulated into a Luré system sector bounded in $[-1, 1]$ and then controller and observer matrices are derived using Lyapunov theory and linear matrix inequalities. Two steps that may introduce conservatism in this approach are (1) reformulating the DANN as a Luré, i.e., going from formulation (5) to (8), and (2) implementing the S-procedure to obtain, for example, (30) from (21) and (29). Although these steps may introduce conservatism, they have been applied to derive stability results in the literature and have given quantitatively useful results in our case studies. The theoretical results were applied to a numerical example and a case study for the control of a highly nonlinear process in which two base flowrates are used to control the pH of two tanks in series. The LMI-based controls had improved closed-loop performance compared to PI controllers in both case studies.

References

Anon (2010). *Theory and practice of pH measurement, no. PN 44-6033/rev. D*. Emerson.

Bland, R. G., Goldfarb, D., & Todd, M. J. (1981). The ellipsoid method: A survey. *Operations Research*, *29*(6), 1039–1091.

Boyd, S., El Ghaoui, L., Feron, E., & Balakrishnan, V. (1994). *Linear matrix inequalities in system and control theory*. SIAM.

Chiang, R. (2019). *MATLAB robust control toolbox*. Natick, MA: MathWorks.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, *2*, 303–314.

Diwadkar, A., Dasgupta, S., & Vaidya, U. (2015). Control of systems in Lure form over erasure channels. *International Journal of Robust and Nonlinear Control*, *25*, 2787–2802.

Drummond, R., Valmorbida, G., & Duncan, S. R. (2018). Generalized absolute stability using Lyapunov functions with relaxed positivity conditions. *IEEE Control Systems Letters*, *2*(2), 207–212.

Faanes, A., & Skogestad, S. (2004). pH-neutralization: Integrated process and control design. *Computers & Chemical Engineering*, *28*, 1475–1487.

Fazlyab, M., Morari, M., & Pappas, G. J. (2020). Safety verification and robustness analysis of neural networks via quadratic constraints and semidefinite programming. arXiv:1903.01287v2.

Funahashi, K. (1989). On the approximate realization of continuous mappings by neural networks. *Neural Networks*, *2*, 183–192.

Gonzaga, C. C., Jungers, M., & Daafouz, J. (2012). Stability analysis of discrete-time Lur'e systems. *Automatica*, *48*(9), 2277–2283.

Gusev, S. V., & Likhtarnikov, A. L. (2006). Kalman-Popov-Yakubovich lemma and the S-procedure: A historical essay. *Automation and Remote Control*, *67*(11), 1768–1810.

Gustafsson, T. K., Skrifvars, B. O., Sandström, K. V., & Waller, K. V. (1995). Modeling of pH for control. *Industrial and Engineering Chemistry Research*, *34*, 820–827.

Hahn, W. (1958). Über die Anwendung der Methode von Ljapunov auf Differenzengleichungen. *Mathematische Annalen*, *136*, 430–441.

Himmelblau, D. M. (2008). Accounts of experiences in the application of artificial neural networks in chemical engineering. *Industrial and Engineering Chemistry Research*, *47*, 5782–5796.
Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, *79*, 2554–2558.
Khalil, H. (2002). *Nonlinear systems* (3rd ed.). Upper Saddle River, NJ, USA: Prentice-Hall.
Kim, K.-K. K., & Braatz, R. D. (2014). Observer-based output feedback control of discrete-time Luré systems with sector-bounded slope-restricted nonlinearities. *International Journal of Robust and Nonlinear Control*, *24*, 2458–2472.
Kim, K.-K. K., Ríos-Patrón, E., & Braatz, R. D. (2018). Standard representation and unified stability analysis for dynamic artificial neural network models. *Neural Networks*, *98*, 251–262.
Legat, B. (2020). *Set programming: Theory and computation* (Ph.D. thesis), Université Catholique de Louvain.
Levin, A. U., & Narendra, K. S. (1995). Recursive identification using feedforward neural networks. *International Journal of Control*, *61*(3), 533–547.
Luré, A. I., & Postnikov, V. N. (1944). On the theory of stability of control systems. *Journal of Applied Mathematics and Mechanics*, *8*(3), 246–248.
Michel, A. N., Farrell, J. A., & Porod, W. (1989). Qualitative analysis of neural networks. *IEEE Transactions on Circuits and Systems*, *36*(2), 229–243.
Monteiro, R., & Todd, M. (2000). Path-following methods. In *Handbook of semidefinite programming* (pp. 267–306). Springer.
Monteiro, R. D. C., & Zanjácomo, P. (1999). Implementation of primal-dual methods for semidefinite programming based on Monteiro and tsuchiya Newton directions and their variants. *Optimization Methods & Software*, *11*(1–4), 91–140.
Morari, M., & Zafiriou, E. (1989). *Robust process control*. Pistcataway, NJ: Prentice Hall.
Nesterov, Y., & Nemirovskii, A. (1994). *Interior-point polynomial algorithms in convex programming*. SIAM.
Nguyen, H. H., Matschek, J., Zieger, T., Savchenko, A., Noroozi, N., & Findeisen, R. (2020). Towards nominal stability certification of deep learning-based controllers. In *American control conference* (pp. 3886–3891).
Nikolakopoulou, A., Hong, M. S., & Braatz, R. D. (2020). Feedback control of dynamic artificial neural networks using linear matrix inequalities. In *IEEE conference on decision and control* (pp. 2210–2215).
Nikolakopoulou, A., Hong, M. S., & Braatz, R. D. (2021). Output feedback control and estimation of dynamic neural networks using linear matrix inequalities. In *American control conference* (pp. 2613–2618).
Park, P. (2002). Stability criteria of sector- and slope-restricted Lur'e systems. *IEEE Transactions on Automatic Control*, *47*(2), 308–313.
Pearson, R. K. (1995). Nonlinear input/output modelling. *Journal of Process Control*, *5*(4), 197–211.
Riggs, J. B., & Rhinehart, R. R. (1990). Method for pH control. U.S. Patent 4, 940, 551.
Ríos-Patrón, E., & Braatz, R. D. (1998). Global stability analysis for discrete-time nonlinear systems. In *American control conference* (pp. 338–342).
Suykens, J. A. K., Curran, P. F., & Chua, L. O. (1999). Robust synthesis for master-slave synchronization of Lur'e systems. *IEEE Transactions on Circuits and Systems–I: Fundamental Theory and Applications*, *46*(7), 841–850.
Suykens, J. A. K., De Moor, B. L. R., & Vandewalle, J. (1995). Nonlinear system identification using neural state space models, applicable to robust control design. *International Journal of Control*, *62*(1), 129–152.

Anastasia Nikolakopoulou was a graduate student at the group of Prof. Richard D. Braatz at the Massachusetts Institute of Technology (MIT). She received a Diploma from the National Technical University of Athens and an M.S. in Chemical Engineering Practice and a Ph.D. from MIT. Her research is in the optimization and control of modular manufacturing systems, state estimation and control of nonlinear dynamical systems, and model predictive control of distributed parameter systems. In 2021 Anastasia moved to an industrial position in the pharmaceutical industry where she is a Scientist in Process Operation and Modeling.

Moo Sun Hong is a postdoctoral researcher in the group of Prof. Richard D. Braatz at the Massachusetts Institute of Technology (MIT). He received an M.S. in Chemical Engineering Practice and Ph.D. from MIT. His research is in model-based design and control and its application to advanced biopharmaceutical manufacturing systems. Honors include the AIChE PD2M Award for Excellence in Integrated QbD Practice, the AIChE Separations Division Graduate Student Research Award, and the Jefferson W. Tester Award from the MIT Chemical Engineering Practice School.

Richard D. Braatz is the Edwin R. Gilliland Professor at the Massachusetts Institute of Technology (MIT) where he does research in control theory and its applications to advanced manufacturing. He received an MS and Ph.D. from the California Institute of Technology and was a Professor at the University of Illinois at Urbana-Champaign and a Visiting Scholar at Harvard University before moving to MIT. Recognitions include the AACC Donald P. Eckman Award and the IEEE CSS Antonio Ruberti Young Researcher Prize. He is a Fellow of IEEE and IFAC and a member of the U.S. National Academy of Engineering.