



Brief Paper

Regional stability conditions for recurrent neural network-based control systems[☆]Alessio La Bella^{a,*}, Marcello Farina^a, William D'Amico^a, Luca Zaccarian^{b,c}^a Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133, Milano, Italy^b CNRS, LAAS, Université de Toulouse, 31400 Toulouse, France^c Dipartimento di Ingegneria Industriale, University of Trento, 38122 Trento, Italy

ARTICLE INFO

Article history:

Received 9 February 2024
 Received in revised form 25 November 2024
 Accepted 8 December 2024
 Available online 24 January 2025

Keywords:

Recurrent neural networks
 Lyapunov stability
 Regional sector condition

ABSTRACT

In this paper we propose novel global and regional stability analysis conditions based on linear matrix inequalities for a general class of recurrent neural networks. These conditions can be also used for state-feedback control design and a suitable optimization problem enforcing \mathcal{H}_2 norm minimization properties is defined. The theoretical results are corroborated by numerical simulations, showing the advantages and limitations of the methods presented herein.

© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Considering the growing availability of data in many engineering fields, there has been an increasing interest on neural networks (NNs) and their properties in the last decades, in particular considering recurrent NNs (RNNs). This is also true in the systems and control realm, where the latter have been subject of intense research in view of their useful features (Lanzetti et al., 2019; Sontag, 1992). In particular, RNNs can be reliably used for modeling complex and nonlinear dynamical systems, given their stateful nature that allows them to store memory of past data, as well as for data-based control design (Bonassi et al., 2022). The so-called *indirect design methods*, for instance, consist of the design of suitable controllers based on models learned from data and, in this context, RNN models are effective as they provide reliable approximations of nonlinear system dynamics. This approach has been adopted for the design of model predictive controllers, e.g., in Bugliari Armenio et al. (2019), Lanzetti et al. (2019). In this context, it is relevant to study systems-theoretic properties of RNNs, e.g., observability,

controllability, and stability of motion. These properties are fundamental, e.g., for characterizing RNN-based models and the ensuing controller designs. Different analysis conditions guaranteeing stability-related properties have been proposed. They include global asymptotic stability (GAS), input-to-state stability (ISS), and incremental ISS (δ ISS) (Bayer et al., 2013; Jiang & Wang, 2001). Sufficient conditions ensuring GAS are derived for RNNs in Stipanović et al. (2021). Alternative contraction properties are characterized in Buehner and Young (2006) for echo state networks (ESNs) and in Revay et al. (2020) for more general RNNs. Sufficient analysis conditions guaranteeing δ ISS are studied in Bonassi et al. (2021, 2023) for long short-term memory networks (LSTMs) and gated recurrent units (GRUs), respectively. Note that δ ISS is a stronger stability property, as it implies ISS and GAS. Thus, in D'Amico et al. (2024) sufficient conditions ensuring δ ISS are proposed for a class of discrete-time nonlinear systems, which includes different common RNN classes, e.g., ESNs and neural nonlinear autoregressive networks with exogenous variables (NNARX). Design conditions for NN-based controllers have been proposed in some relevant recent works (Barabanov & Prokhorov, 2002; Liu et al., 2021). Design conditions for FFNN controllers are also provided in Vance and Jagannathan (2008), considering specific classes of second-order nonlinear systems. More recently, event-triggered FFNN controllers have been proposed in de Souza, Tarbouriech et al. (2023) for linear systems and in de Souza, Girard et al. (2023) for perturbed ones, providing conditions to ensure local stability properties for the closed-loop system. In this context, Yin et al. (2021) developed conditions for stability analysis of closed-loop systems composed of a pre-designed FFNN-based controller and perturbed linear plant. The main limitations of results listed above, except for de Souza,

[☆] The research activity has been carried out within the project "SafeTy Aware Reinforcement Learning for robotic inspection" (STARLIT) funded by the Italian MUR, CUP n. E53D23001130006, number 2022ZE9J9J and by the ANR via grant OLYMPIA, Italy, number ANR-23-CE48-0006. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Laurentiu Hetel under the direction of Editor Sophie Tarbouriech.

* Corresponding author.

E-mail addresses: alessio.labella@polimi.it (A. La Bella), marcello.farina@polimi.it (M. Farina), william.damico@polimi.it (W. D'Amico), luca.zaccarian@laas.fr (L. Zaccarian).

Tarbouriech et al. (2023), de Souza, Girard et al. (2023), and Yin et al. (2021), is that they provide global stability properties. However, global stability conditions could be not always admissible. This may be due, e.g., to the RNNs nonlinear dynamics or to the boundedness of some state and input variables in a closed-loop scenario where the plant exhibits unstable dynamics. Similarly to what is done in Massimetti et al. (2009) and in Da Silva and Tarbouriech (2001) in the case of linear plants with input saturation, it is necessary to develop regional stability conditions, associated with an estimate of the basin of attraction of the considered equilibrium.

In view of the above discussion, this paper proposes novel global and local stability analysis conditions based on linear matrix inequalities (LMIs) for a general class of RNNs. These properties can be established thanks to the peculiarities of sigmoidal nonlinearities typically used in the RNN models. We also show that these conditions can be enforced in state-feedback control design. The generalization to more complex control systems (including, e.g., state observers, integrators, etc.) is rather straightforward and can be done following the lead of the present paper. The theoretical results are corroborated by numerical simulations, showing the advantages and limitations of the methods presented herein.

The problem and the RNN model are introduced in Section 2. In Section 3, a global stability condition for RNN models is derived and then, given its limitation, two different regional stability conditions are introduced. The derived conditions are then exploited for control design in Section 4. Simulation results are illustrated in Section 5 and conclusions are drawn in Section 6.

Notation. Let \mathbb{R} denote the set of real numbers, $\mathbb{R}_{\geq 0}$ the set of positive or null real numbers including zero, and $\mathbb{R}_{> 0}$ the set of strictly positive real numbers. Given a matrix $A \in \mathbb{R}^{n \times n}$, its transpose is A^T , whereas its eigenvalues are $\lambda_i(A)$, with $i = 1 \dots, n$. The i th entry of a vector v is indicated as v_i . The entry in the i th row and j th column of a matrix A is denoted a_{ij} , whereas the i th row and the j th column vectors of A are indicated as $A_{i \cdot}$ and $A_{\cdot j}$, respectively. Given a matrix P , we use $P \geq 0$, $P > 0$, $P \leq 0$, and $P < 0$ to indicate that it is positive semidefinite, positive definite, negative semidefinite, and negative definite, respectively. $\lambda_{\min}(P)$ and $\lambda_{\max}(P)$ denote the minimum and maximum eigenvalues of a symmetric matrix P , respectively. $0_{n,m}$ denotes a zero matrix with n rows and m columns and I_n is the identity matrix of dimension n . Given a vector v , $\|v\| = \sqrt{v^T v}$ denotes the 2-norm of a column vector v and $\|v\|_Q = \sqrt{v^T Q v}$ denotes the weighted Euclidean norm of v , where Q is a positive definite matrix. Given a vector $v \in \mathbb{R}^n$, we define its infinity norm as $\|v\|_\infty = \max_{i=1, \dots, n}(|v_i|)$. A matrix $A \in \mathbb{R}^{n \times n}$ is said to be Schur if $\|\lambda_i(A)\| < 1$, for each $i = 1, \dots, n$. The set of positive definite symmetric real matrices is denoted as $\mathbb{S}_{> 0}^n = \{A \in \mathbb{R}^{n \times n} : A = A^T \wedge A > 0\}$, whereas the set of diagonal positive definite real matrices is denoted as $\mathbb{D}_{> 0}^n = \{A \in \mathbb{S}_{> 0}^n : a_{ij} = 0 \ \forall i, j \in [1, \dots, n] \wedge i \neq j\}$.

2. Problem statement

This work is concerned with stability analysis and control design for a rather general class of Recurrent Neural Networks (RNNs) representing the plant model. We assume that the following discrete-time model, derived from a batch of input-output data (u, z_m) , accurately represents the plant dynamics,

$$x^+ = A_o x + B_u u + B_\sigma \sigma(C_o x + D_u u) \quad (1a)$$

$$z_m = C_m x, \quad (1b)$$

$x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^m$ the input vector, $z_m \in \mathbb{R}^p$ the measurements vector, $y \mapsto \sigma(y) = [\sigma_1(y_1) \dots \sigma_\nu(y_\nu)]^T$ a vector

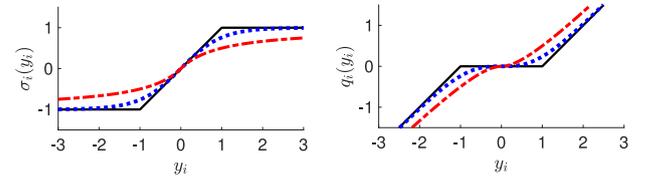


Fig. 1. Left plot: Examples of functions σ_i satisfying Assumption 1: $\sigma_i(y_i) = \text{sat}(y_i)$ (solid black line), $\sigma_i(y_i) = \tanh(y_i)$ (dotted blue line) and $\sigma_i(y_i) = y_i/(1 + |y_i|)$ (dashed red line); Right plot: corresponding functions $q_i(y_i)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of decentralized functions $\sigma_i : \mathbb{R} \rightarrow \mathbb{R}$ better specified below, $A_o \in \mathbb{R}^{n \times n}$, $B_u \in \mathbb{R}^{n \times m}$, $B_\sigma \in \mathbb{R}^{n \times \nu}$, $C_o \in \mathbb{R}^{\nu \times n}$, and $D_u \in \mathbb{R}^{\nu \times m}$. The following assumption is stated.

Assumption 1. Each component $\sigma_i : \mathbb{R} \rightarrow \mathbb{R}$ of function $\sigma : \mathbb{R}^\nu \rightarrow \mathbb{R}^\nu$ in (1), i.e., with $i \in \{1, \dots, \nu\}$, is a sigmoid function, namely, σ_i is a monotone increasing function, globally Lipschitz continuous with unitary Lipschitz constant, and such that $\sigma_i(0) = 0$, $\sigma_i'(0) = \left. \frac{d\sigma_i(y_i)}{dy_i} \right|_{y_i=0} = 1$, and $\sigma_i(y_i) \in [-1, 1]$ for all $y_i \in \mathbb{R}$.

Assumption 1 is satisfied by different sigmoid functions, see Fig. 1. For instance, model (1) under Assumption 1 represents generic classes of RNNs, e.g., Echo State Networks (ESNs) and shallow Neural Nonlinear AutoRegressive eXogenous (NNARX) networks. The reader is referred to D'Amico et al. (2024, Section II) for more details on how the mentioned RNN families can be formulated as in (1). Note that σ can comprise different scalar components σ_i , provided that they all respect Assumption 1.

The purpose of this work is twofold. One goal is to establish LMI-based analysis conditions guaranteeing asymptotic stability of the origin for (1) under linear state feedback $u = Kx$. Both global and local stability conditions will be proposed, with a possibly maximized ellipsoidal estimate of the basin of attraction. Then, the analysis conditions will be exploited for the design of a state feedback gain K , ensuring closed-loop stability and suitable performance. Note that the control law has state-feedback form just for the sake of simplicity, but a suitable observer should be used to provide state estimates through input-output data (this aspect will be later discussed in Remark 3); in fact, as usual in case of RNNs, the states of model (1) may not correspond to real plant physical and measurable variables. The resulting closed-loop system is

$$\begin{cases} x^+ = Ax + B(y - \sigma(y)) \\ \quad = Ax + Bq(y), \\ y = Cx, \end{cases} \quad (2)$$

where $C = C_o + D_u K$, $A = A_o + B_u K + B_\sigma C$, $B = -B_\sigma$, and $q : \mathbb{R}^\nu \rightarrow \mathbb{R}^\nu$ is a vector of scalar functions $y \mapsto q(y) = [q_1(y_1) \dots q_\nu(y_\nu)]^T$ with $q_i(y_i) = y_i - \sigma_i(y_i)$, for any $y_i \in \mathbb{R}$ and for each $i \in \{1, \dots, \nu\}$. Fig. 1 depicts $\sigma_i(\cdot)$ and $q_i(\cdot)$ for different functions respecting Assumption 1, comparing them with two notable functions: the saturation $y \mapsto \text{sat}(y) = [\text{sat}_1(y_1) \dots \text{sat}_\nu(y_\nu)]$ and the deadzone $y \mapsto \text{dz}(y) = y - \text{sat}(y)$, respectively, whose components are defined as $\text{sat}_i(y_i) = \max(-1, \min(y_i, 1))$ and $\text{dz}_i(y_i) = y_i - \text{sat}_i(y_i)$. Note that, under Assumption 1, since $\sigma(0) = 0$, the origin is an equilibrium for (2). Thus, analysis and design conditions are here studied for K ensuring exponential stability of the origin for the closed-loop system in (2).

Remark 1. The following results can be extended to plant models (1) regulated by dynamic and nonlinear controllers provided that the control law can be expressed as a linear function of the states of the overall control system, as done in D'Amico et al. (2024).

3. Stability analysis

3.1. Global stability analysis via sector conditions

Considering global exponential stability of system (1), a wide literature is available for the specific case $\sigma(\cdot) = \text{sat}(\cdot)$, where the sector characterization of this non-linearity is exploited (Da Silva & Tarbouriech, 2001; Massimetti et al., 2009). We show here that these conditions can be applied for any σ under Assumption 1.

Lemma 1. *Let function $\sigma : \mathbb{R}^v \rightarrow \mathbb{R}^v$ satisfy Assumption 1 and consider $q : \mathbb{R}^v \rightarrow \mathbb{R}^v$ defined as $y \mapsto q(y) = y - \sigma(y)$. Then, for any $y \in \mathbb{R}^v$ and $W \in \mathbb{D}_{>0}^v$, it holds that*

$$q(y)^\top W \sigma(y) \geq 0. \quad (3)$$

The proof of Lemma 1 is reported in La Bella et al. (2024) and its validity is notable from Fig. 1. In view of this result, the following stability condition can be stated.

Proposition 1. *Under Assumption 1, the origin is a globally exponentially stable equilibrium for system (2) if there exist $S \in \mathbb{S}_{>0}^n$ and $U \in \mathbb{D}_{>0}^v$, satisfying the LMI*

$$\begin{bmatrix} S & -SC^\top & SA^\top \\ -CS & 2U & UB^\top \\ AS & BU & S \end{bmatrix} \succ 0. \quad (4)$$

The proof of Proposition 1 derives directly from (Massimetti et al., 2009), exploiting the fact that (3) holds not only for the saturation case but for any function σ satisfying Assumption 1. Nevertheless, the global stability condition in Proposition 1 can be too conservative in some cases, as characterized in the next lemma.

Lemma 2. *There exists a solution to (4) only if both A and $A_0 + B_u K$ are Schur with a common quadratic Lyapunov function $V(x) = x^\top S^{-1}x$.*

The proof Lemma 2 is straightforward from the application of the Schur complement. An extensive proof is reported in La Bella et al. (2024). Note that $A_0 + B_u K$ is not Schur in closed loop schemes including an integral action, therefore condition (4) can never hold in those cases, as discussed in D'Amico et al. (2024). The intrinsic limitation highlighted by Lemma 2 motivates the need for regional stability conditions for system (2), associated with an estimate of the basin of attraction of the stable equilibrium.

A sufficient condition guaranteeing regional stability of system (2) is established in Massimetti et al. (2009) for the saturation case, i.e., if $\sigma(\cdot) = \text{sat}(\cdot)$ and, as a consequence, $q(\cdot) = \text{dz}(\cdot)$. The idea is to guarantee a Lyapunov decrease in a polytopic region characterized by $\|Hx\|_\infty \leq 1$, where $H \in \mathbb{R}^{v \times n}$ is a further degree of freedom, used to enlarge, as much as possible, the estimate of the basin of attraction estimate of the origin. The key result for proving such a regional property derives from the fact the following sector characterization holds

$$\text{dz}(y)^\top W (\text{sat}(y) + Hx) \geq 0, \quad (5)$$

$\forall y \in \mathbb{R}^v$ and $\forall x \in \mathcal{L}(H) = \{x \in \mathbb{R}^n : \|Hx\|_\infty \leq 1\}$. Nevertheless, condition (5) holds specifically for the saturation case and not for any function σ satisfying Assumption 1, thus implying that the regional stability results in Massimetti et al. (2009) cannot be extended to our case. This motivates the need of establishing novel regional stability conditions for a generic system (2) satisfying Assumption 1, representative of many RNN families. Thus, two novel regional stability conditions are proposed in the next two sections. The first one is associated with a single LMI condition, whereas the second one requires solving an iterative LMI-based algorithm. The two conditions exhibit different advantages and uses, as illustrated in Section 5.

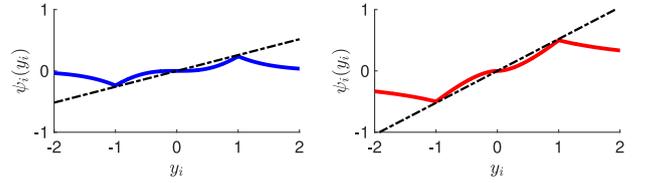


Fig. 2. Function $\psi_i(y_i)$ (solid line) and the sector bound $\theta_i y_i$ (dashed black line). Left plot: $\sigma_i = \tanh(y_i)$, with $\theta_i = 0.2384$; Right plot: $\sigma_i(y_i) = y_i / (1 + |y_i|)$, with $\theta_i = 0.52$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2. Regional stability analysis via auxiliary function

Consider system (2) under Assumption 1 and define the auxiliary function

$$\psi(y) = q(y) - \text{dz}(y) = \text{sat}(y) - \sigma(y). \quad (6)$$

Each element of the function ψ corresponds to the difference between the saturation and the function σ . Introducing the function $\psi(y)$ has a twofold advantage. On one hand, it allows one to exploit the regional sector condition (5) by simply replacing $\text{dz}(y)$ by $q(y) - \psi(y)$ and by recalling that $\text{sat}(y) = y - \text{dz}(y)$. On the other hand, the function ψ allows establishing an additional sector condition, as evident from the next lemma.

Lemma 3. *Under Assumption 1, function $y \mapsto \psi(y)$ in (6) belongs to the sector $[0, \Theta]$ for some $\Theta \in \mathbb{D}_{>0}^v$ where $0 \leq \Theta < 1$. Equivalently, for any $Y \in \mathbb{D}_{>0}^v$, it holds that*

$$\psi(y)^\top Y (\Theta y - \psi(y)) \geq 0, \quad \forall y \in \mathbb{R}^v. \quad (7)$$

The technical proof of Lemma 3 is reported in La Bella et al. (2024). The validity of condition (7) is apparent from Fig. 2, which reports the scalar component $\psi_i(y_i)$ for two different sigmoid functions $\sigma_i(y_i)$, together with the corresponding minimum sector bound $\theta_i y_i$. Note that, for a given Θ , the validity of Lemma 3 can be numerically certified through the COQ proof assistant (Chlipala, 2022). Considering the functions depicted in Fig. 2, the corresponding COQ proof code is available at La Bella (2024). Before stating the main result, note that, in view of (6), the sector condition (5) can be rewritten as

$$(q(y) - \psi(y))^\top W (y - q(y) + \psi(y) + Hx) \geq 0, \quad (8)$$

which, notably, holds for all $x \in \mathcal{L}(H)$ and all $y \in \mathbb{R}^v$. Note that the sector condition (5) only holds for the $\text{sat}(\cdot)$ nonlinearity and does not hold for a generic σ satisfying Assumption 1. Nevertheless, combining (5) with the sector condition (7), which bounds the difference between the saturation and the sigmoid function, i.e., $\psi(\cdot)$, the novel regional stability condition (8) is enjoyed by system (2) under Assumption 1.

Theorem 1. *The origin is a locally exponentially stable equilibrium for system (2) under Assumption 1 if there exist $S \in \mathbb{S}_{>0}^n$, $U \in \mathbb{D}_{>0}^v$, $R \in \mathbb{D}_{>0}^v$, and $L \in \mathbb{R}^{v \times n}$, satisfying the following linear matrix inequality system*

$$\begin{bmatrix} S & -L^\top - SC^\top & -SC^\top \Theta & SA^\top \\ -L - CS & 2U & 0 & UB^\top \\ -\Theta CS & 0 & 2R & RB^\top \\ AS & BU & BR & S \end{bmatrix} \succ 0, \quad (9a)$$

$$\begin{bmatrix} S & L_{i,:}^\top \\ L_{i,:} & 1 \end{bmatrix} \succeq 0, \quad \forall i \in \{1, \dots, v\}, \quad (9b)$$

where Θ is defined such that (7) holds. Moreover, the set $\mathcal{E}(S) = \{x : x^\top S^{-1}x \leq 1\}$ satisfies $\mathcal{E}(S) \subseteq \mathcal{L}(H)$, with $H = LS^{-1}$, and it is a positively invariant set contained in the basin of attraction of the origin.

Proof of Theorem 1. We first show that (9b) implies $\mathcal{E}(S) \subseteq \mathcal{L}(H)$. Considering $P = S^{-1}$ with $S \in \mathbb{S}_{>0}^n$, (9b) can be rewritten as $\begin{bmatrix} S & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} P & H_{i,:}^\top \\ H_{i,:} & 1 \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & 1 \end{bmatrix} \geq 0$, which implies, by means of Schur complement, that

$$x^\top H_{i,:}^\top H_{i,:} x \leq x^\top P x \leq 1, \quad (10)$$

for any $x \in \mathcal{E}(S)$ and for all $i \in \{1, \dots, \nu\}$. Given the definition of $\mathcal{L}(H)$, then (10) implies that $\mathcal{E}(S) \subseteq \mathcal{L}(H)$. Consider now the candidate Lyapunov function $V(x) = x^\top P x$ for system (2). To prove the local stability of the origin in $\mathcal{E}(S) = \mathcal{E}(P^{-1})$, it is sufficient to show that $\Delta V(x) = V(x^+) - V(x)$, for any $x \in \mathcal{E}(S) \setminus \{0\}$, is

$$\Delta V(x) = \tilde{x}^\top \left(\begin{bmatrix} A^\top \\ B^\top \\ 0 \end{bmatrix} P \begin{bmatrix} A & B & 0 \end{bmatrix} - \begin{bmatrix} P & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) \tilde{x} < 0, \quad (11)$$

where $\tilde{x} = [x^\top \ q(Cx)^\top \ \psi(Cx)^\top]^\top$. Given that $\mathcal{E}(S) \subseteq \mathcal{L}(H)$, the sector condition (8) can be exploited to prove the validity of (11). To this end, (8) is rewritten as

$$\tilde{x}^\top \begin{bmatrix} 0 & \Phi^\top & -\Phi^\top \\ \Phi & -2W & 2W \\ -\Phi & 2W & -2W \end{bmatrix} \tilde{x} \geq 0, \quad \forall x \in \mathcal{L}(H), \quad (12)$$

where $\Phi = W(C+H)$ is introduced for the sake of compactness. Analogously, condition (7) can be rewritten as

$$\tilde{x}^\top \begin{bmatrix} 0 & 0 & C^\top \Theta Y \\ 0 & 0 & 0 \\ Y \Theta C & 0 & -2Y \end{bmatrix} \tilde{x} \geq 0. \quad (13)$$

Thus, by leveraging the S-procedure (Boyd et al., 1994), it is possible to guarantee (11) under (12) and (13) if $\exists \tau_1, \tau_2 > 0$ such that

$$\begin{bmatrix} A^\top \\ B^\top \\ 0 \end{bmatrix} P \begin{bmatrix} A & B & 0 \end{bmatrix} - \begin{bmatrix} P & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \tau_1 \begin{bmatrix} 0 & \Phi^\top & -\Phi^\top \\ \Phi & -2W & 2W \\ -\Phi & 2W & -2W \end{bmatrix} + \tau_2 \begin{bmatrix} 0 & 0 & C^\top \Theta Y \\ 0 & 0 & 0 \\ Y \Theta C & 0 & -2Y \end{bmatrix} < 0. \quad (14)$$

Now, introducing $\tilde{W} = \tau_1 W$, $\tilde{Y} = \tau_2 Y$, and $\tilde{\Phi} = \tau_1 \Phi = \tilde{W}(C+H)$, by Schur complement, (14) can be written as

$$\begin{bmatrix} P & -\tilde{\Phi}^\top & \tilde{\Phi}^\top - C^\top \Theta \tilde{Y} & A^\top \\ -\tilde{\Phi} & 2\tilde{W} & -2\tilde{W} & B^\top \\ \tilde{\Phi} - \tilde{Y} \Theta C & -2\tilde{W} & 2\tilde{W} + 2\tilde{Y} & 0 \\ A & B & 0 & P^{-1} \end{bmatrix} > 0. \quad (15)$$

Finally, consider $U = \tilde{W}^{-1}$, $R = \tilde{Y}^{-1}$, and the matrix $M_1 = \begin{bmatrix} S & 0 & 0 & 0 \\ 0 & U & 0 & 0 \\ 0 & R & R & 0 \\ 0 & 0 & 0 & I_n \end{bmatrix}$. Pre and post-multiplying (15) by M_1 and its transpose, respectively, (9a) is obtained. Thus, if (9) holds, (11) is verified for all $x \in \mathcal{E}(S) \setminus \{0\}$ and, consequently, the origin is locally exponentially stable with a basin of attraction including $\mathcal{E}(S)$. \square

Comparing condition (9a) with (4), it is worth noting that the additional decision variable L (related to H) allows for more degrees of freedom in the LMI feasibility problem. As a result, no condition must be imposed on $A_o + B_u K$ for the feasibility of (9a), thus overcoming the limitations of Lemma 2, as explained in the next lemma, whose proof is reported in La Bella et al. (2024).

Lemma 4. A necessary and sufficient condition for the existence of a small enough $\Theta > 0$ ensuring feasibility of (9) is that A is Schur stable.

The condition provided in Theorem 1 is a powerful tool not only to guarantee local exponential stability of the origin for system (2), but also to maximize an estimate of the basin of

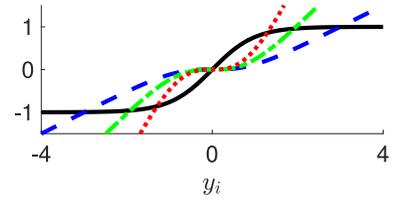


Fig. 3. Function $\sigma_i(y_i) = \tanh(y_i)$ (solid black) and the ensuing function $h_i q_i(y_i)$ for $h_i = 0.5$ (dashed blue), $h_i = 1$ (dashed-dotted green), $h_i = 2$ (dotted red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

attraction. This can be done, for example, by solving the following LMI optimization problem

$$\begin{aligned} \max \quad & \gamma \\ \text{s.t.} \quad & S \in \mathbb{S}_{>0}^n, L \in \mathbb{R}^{\nu \times n}, \\ & U, R \in \mathbb{D}_{>0}^\nu, \gamma \in \mathbb{R} \end{aligned} \quad (16)$$

subject to (9) and $f_{\text{LMI}}(S) \geq \gamma I$,

where $f_{\text{LMI}}(S)$ can be selected by the designer based on the goal. For instance, as discussed in Boyd et al. (1994, Section 5.2), if $f_{\text{LMI}}(S) = \log \det(S)$, the volume of the invariant ellipsoid $\mathcal{E}(S)$ will be maximized. On the other hand, if $f_{\text{LMI}}(S) = S$, the radius of the maximal ball inscribed in $\mathcal{E}(S)$ will be maximized. In both cases the optimization problem is convex. Alternative polytopic inner approximation can be also used. As a final remark, note that the use of two joint sector conditions, i.e., (5) and (8), can make (9) computationally heavy, as it will be evident in Section 5. For this reason, a less complex, but less computationally attractive, regional stability condition for system (2) is proposed in the following section.

3.3. Regional stability analysis via sector narrowing

An alternative parametric sector condition enjoyed by $\sigma(y)$ and $q(y) = y - \sigma(y)$ can be provided, leading to the development of a different regional stability condition, described in the following lemma, whose proof is reported in La Bella et al. (2024).

Lemma 5. Under Assumption 1, there exists a monotone decreasing function $\tilde{y}_i : \mathbb{R} \rightarrow \mathbb{R}$, such that, for any $h_i > 0$, it holds that

$$y_i(\sigma_i(y_i) - h_i q_i(y_i)) \geq 0, \quad \forall y_i \in \mathcal{Y}_i(h_i), \quad (17)$$

where $\mathcal{Y}_i(h_i) = [-\tilde{y}_i(h_i), \tilde{y}_i(h_i)]$.

The validity of condition (17) is also illustrated in Fig. 3, where the scalar function $\sigma_i(y_i)$ and the ensuing lower bound $h_i q_i(y_i)$ are depicted for a specific sigmoid function and three specific values of h_i . Introducing $\mathcal{Y}(H) = \mathcal{Y}_1(h_1) \times \dots \times \mathcal{Y}_\nu(h_\nu)$, where $H = \text{diag}(h_1, \dots, h_\nu)$, inequality (17) in Lemma 5 can be also expressed in matrix form. As a consequence, for any $H \in \mathbb{D}_{>0}^\nu$ and any $W \in \mathbb{D}_{>0}^\nu$, there exists $\mathcal{Y}(H)$ such that

$$q(y)^\top W(\sigma(y) - Hq(y)) \geq 0, \quad \forall y \in \mathcal{Y}(H). \quad (18)$$

In general, an explicit expression of $\mathcal{Y}(H)$ cannot be computed. However, a lower bound of $\mathcal{Y}(H)$ can be numerically computed based on the choice of H and on the specific sigmoid function σ . To do that, for a given $H = \text{diag}(h_1, \dots, h_\nu)$, it is sufficient to numerically compute the values of $y_i \neq 0$ such that condition (17) holds with an equality, for each $i = 1, \dots, \nu$. Considering σ_i to be an odd function for simplicity and given its sigmoidal structure defined in Assumption 1 (see Fig. 3), two nonzero solutions exist, i.e., $y_i = \tilde{y}_i(h_i)$ and $y_i = -\tilde{y}_i(h_i)$, characterizing the bounds of $\mathcal{Y}_i(h_i)$ for a given h_i . Also in this case, for given a H and a corresponding set $\mathcal{Y}(H)$, the validity of (18) can be numerically certified exploiting the COQ assistant (Chlipala, 2022). The COQ

code validating the sector condition for the considered h_i in Fig. 3 is available at La Bella (2024). Based on the regional sector condition (18), the second main analysis result of this work can be stated.

Theorem 2. *The origin is a locally exponentially stable equilibrium for system (2) under Assumption 1, if there exists $S \in \mathbb{S}_{>0}^n$, $U \in \mathbb{D}_{>0}^v$, and $H = \text{diag}(h_1, \dots, h_v) \in \mathbb{D}_{>0}^v$, such that*

$$\begin{bmatrix} S & -SC^T & SA^T \\ -CS & 2(H+I)U & UB^T \\ AS & BU & S \end{bmatrix} > 0, \quad (19a)$$

$$\begin{bmatrix} S & SC_{i,:}^T \\ C_{i,:}S & \bar{y}_i^2(h_i) \end{bmatrix} \geq 0, \quad \forall i \in \{1, \dots, v\}, \quad (19b)$$

with $\bar{y}_i(h_i)$ defined such that (17) holds. Moreover, the set $\mathcal{E}(S) = \{x : x^T S^{-1} x \leq 1\}$ is a forward invariant set and it is contained in the basin of attraction of the origin. Finally, for all $x \in \mathcal{E}(S)$, it holds that $y = Cx \in \mathcal{Y}(H)$.

Proof of Theorem 2. Following similar steps to the ones before (10) and exploiting the definition of $\mathcal{E}(S)$, with $P = S^{-1} \in \mathbb{S}_{>0}^n$, (19) implies that, for any $x \in \mathcal{E}(S)$, it holds that $\frac{x^T C_{i,:}^T C_{i,:} x}{\bar{y}_i^2(h_i)} \leq x^T P x \leq 1$, for all $i \in \{1, \dots, v\}$. This shows that $y = Cx \in \mathcal{Y}(H)$ for all $x \in \mathcal{E}(S)$. Consider now $V(x) = x^T P x$ as a candidate Lyapunov function. Local exponential stability of the origin in $\mathcal{E}(S) = \mathcal{E}(P^{-1})$ holds if, $\forall x \in \mathcal{E}(S) \setminus \{0\}$,

$$\Delta V(x) = V(x^+) - V(x) =$$

$$[x^T \ q(Cx)^T] \left(\begin{bmatrix} A^T \\ B^T \end{bmatrix} P \begin{bmatrix} A & B \end{bmatrix} - \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} \right) \begin{bmatrix} x \\ q(Cx) \end{bmatrix} < 0. \quad (20)$$

Given $y = Cx \in \mathcal{Y}(H)$ for all $x \in \mathcal{E}(S)$, (18) can be exploited to prove (20), by imposing that $\Delta V(x) + q(Cx)^T W(Cx - q(Cx) - Hq(Cx)) < 0$, for any $x \in \mathcal{E}(S) \setminus \{0\}$. By developing the quadratic form of the latter inequality, it follows that (20) holds if $\begin{bmatrix} A^T \\ B^T \end{bmatrix} P \begin{bmatrix} A & B \end{bmatrix} - \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & C^T W \\ WC & -2W(H+I) \end{bmatrix} < 0$, which, by a Schur complement, holds if, and only if,

$$\begin{bmatrix} P & -C^T W & A^T \\ -WC & 2W(H+I) & B^T \\ A & B & P^{-1} \end{bmatrix} > 0. \quad (21)$$

Finally, define $U = W^{-1}$ and $M_2 = \text{diag}(S, U, I)$. Then, pre and post-multiplying (21) by M_2 , (19) is obtained. Therefore, (19) implies the validity of (20), proving local exponential stability of the origin with basin of attraction containing $\mathcal{E}(S)$. \square

Also in this case, the limitation of Lemma 2 are overcome by means of regional stability analysis. The following lemma, whose proof is reported in La Bella et al. (2024), provides a necessary and sufficient condition for the feasibility of the matrix inequality (19).

Lemma 6. *A necessary and sufficient condition for the feasibility of (19) is that A is Schur stable.*

Following Lemma 6, the next corollary holds, whose proof is available in La Bella et al. (2024).

Corollary 1. *If A is Schur stable, a scalar $\bar{h} > 0$ exists such that (19) admits a feasible solution for any $H \geq \bar{h}I$.*

Since $\bar{y}_i(\cdot)$ in Lemma 5 is a monotone decreasing function, it is expected that large estimates of the basin of attraction are achieved close to small values of H (and of \bar{h}), since these correspond to large regional sectors $\mathcal{Y}(H)$. The value \bar{h} constructed in Corollary 1 may be conservative in terms of size of the estimated

Algorithm 1 Optimization of $\mathcal{E}(S)$ through (19).

Ensure: A in system (2) is Schur stable

- 1: **Set** $i = 0$ and $\Delta h > 0$ to a suitable small parameter
- 2: **Solve** (22) and select $\bar{h} = \lambda_{\max}(H_u U^{-1})$
- 3: **while** $i \leq i_{\max}$ **do**
- 4: **Set** $H = (\bar{h} + i\Delta h)I$
- 5: **Solve** (23), and denote the optimal solution as $S^{(i)}, \gamma^{(i)}$
- 6: **Set** $i = i + 1$
- 7: **end while**
- 8: **Select** $i_{\text{best}} = \underset{i=0, \dots, i_{\max}}{\text{argmax}} \gamma^{(i)}$

The best estimate of the basin of attraction of the origin is $\mathcal{E}(S^{(i_{\text{best}})})$.

basin of attraction. It is therefore suggested to select an optimized H according to the following LMI optimization problem

$$\min_{\substack{S \in \mathbb{S}_{>0}^n, U \in \mathbb{D}_{>0}^v, \\ H_u \in \mathbb{D}_{>0}^v, \gamma_u \in \mathbb{R}}} -\gamma_u \quad (22a)$$

subject to

$$\begin{bmatrix} S & -SC^T & SA^T \\ -CS & 2(H_u + U) & UB^T \\ AS & BU & S \end{bmatrix} > 0, \quad (22b)$$

$$U \geq I, \quad (22c)$$

$$H_u \leq \gamma_u I. \quad (22d)$$

where $H_u = HU$ is introduced. Note that (22c) is imposed (in place of $U > 0$) in view of the fact that (22) is homogeneous in the decision variables (U, H_u, S) . In fact, one may select $\bar{h} = \gamma_u$ in Corollary 1 because any feasible solution to (22) satisfies $H = H_u U^{-1} \leq H_u \leq \gamma_u I$, so that any larger selection of H is a feasible one. More generally, one may obtain a potentially less conservative estimate of the minimum \bar{h} by selecting $\bar{h} = \lambda_{\max}(H_u U^{-1}) \leq \gamma_u$.

As done in Section 3.2, it is possible to maximize the size of the basin of attraction estimate $\mathcal{E}(S)$, introduced in Theorem 2. Given H (possibly selected according to (22)), similarly to (16), the size of $\mathcal{E}(S)$ can be maximized by solving the following optimization problem

$$\max_{\substack{S \in \mathbb{S}_{>0}^n, U \in \mathbb{D}_{>0}^v, \\ \gamma \in \mathbb{R}}} \gamma \quad (23)$$

subject to (19) and $f_{\text{LMI}}(S) \geq \gamma I$.

for different possible functions $f_{\text{LMI}}(S)$, as discussed in Section 3.2. Overall, the estimate $\mathcal{E}(S)$ can be maximized through the procedure described in Algorithm 1, which iteratively adjusts the selection of H such that the largest estimate of the basin of attraction is computed.

In particular, Algorithm 1 is structured as follows. Initially, (22) is solved to find the minimum value of \bar{h} such that (19) is feasible. Then, H is iteratively increased for a fixed number of iterations (i.e., i_{\max}) and the estimate of the basin of attraction is computed at each iteration. Ultimately, the iteration corresponding to the largest estimate of the basin of attraction is selected.

3.4. Choice of the regional stability condition

The regional condition based on the auxiliary function (i.e., (9)) is recommended when wanting a prompt application. Indeed, (9) is associated with a single LMI problem, i.e., (16), whereas when

the sector narrowing conditions (19) are exploited, more cumbersome iterative LMIs must be solved to find the matrix H returning the largest estimate of the basin of attraction (cf. Algorithm 1), which may be time consuming. On the other hand, condition (19) and Algorithm 1 are preferred when the user aims to deal with large-order systems. Indeed, as shown in Section 5, the condition based on the auxiliary function (i.e., (9)) leads to numerical problems in case of high-dimensional systems, while the one based on sector narrowing (i.e., (19)) is still feasible, corroborating the potential of the latter condition. On the other hand, results in Section 5 show that, when both regional stability conditions are feasible, the one based on the auxiliary function, i.e., (9), returns a larger estimate of the basin of attraction. Finally, it is worth noting that, for both conditions, the dimension of the considered regional sectors, i.e., $\mathcal{L}(H)$ for the LMI problem (16) and $\mathcal{Y}(H)$ for Algorithm 1, is optimized in order to maximize the size of the estimate of the basin of attraction. This is different from Yin et al. (2021), where regional sector bounds are a-priori fixed and then local stability properties are assessed.

Remark 2. The robustness of the stability properties of the equilibrium of the control system with respect to possible uncertainties is a key feature enabling the controller to be efficiently applied. In our case, robustness is guaranteed, besides by the inherent robustness of the stability of Lyapunov stable equilibria (see Goebel et al. (2012, Chapter 7)), by the fact that our conditions are not defined for the specific model (1) solely, but for all nonlinearities $\sigma(\cdot)$ fulfilling the related sector conditions. Note that, while there may not be an exact correspondence between model (1) and the real plant, the estimate basin of attraction gives a measure of the range of validity of the regional stability properties. Even more so, through the output map (1b) it is possible to define a correspondence between such positively invariant set and the region of validity of our controller for the plant quantities; if necessary, the computation of S can be done by maximizing, at the same time, the size of its projection onto the output space through a suitable choice of function $f_{\text{LMI}}(\cdot)$.

Remark 3. As discussed in Section 2, the considered state feedback control scheme may necessitate the design of a suitable state observer, as typical for RNN models. The scheme would then be perturbed by a vanishing estimation error, which does not invalidate the established stability properties. Indeed, one may follow similar constructions to those of δ ISS observers (D'Amico et al., 2023a), which can be easily designed for (2) considering that σ , under Assumption 1, is incrementally sector bounded.

4. Controller design

As previously discussed, the conditions presented in Section 3 are quite effective for analyzing the regional stability of the origin for system (2), but they can also be an efficient tool for the design of stabilizing controllers. Hence, the objective addressed in this section is to show how to design a state-feedback control law for system (2), i.e., $u_k = Kx_k$, such that: (i) regional (or global) asymptotic stability of the origin is ensured, (ii) an estimate of the basin of attraction is determined, and (iii) desired closed-loop performances are enforced. To this end, introduce $F = A_o + B_o C_o$, $G = B_u + B_\sigma D_u$. It is worth noting that, for achieving global stability of the closed loop, it is necessary that a K exists such that both $A_o + B_u K$ and $F + GK$ are Schur with a common Lyapunov function (in view of Lemma 2). On the other hand, for regional stability is enough that K exists such that $F + GK$ is Schur (in view of Lemmas 4 and 6). Considering the mentioned objectives, Theorem 1 or 2 can be exploited. Defining $J = KS$, condition (9a)

in Theorem 1 can be written as

$$\begin{bmatrix} S & -L - \Omega^\top & -\Omega^\top \Theta & \Lambda^\top \\ -L - \Omega & 2U & 0 & UB^\top \\ -\Theta \Omega & 0 & 2R & RB^\top \\ \Lambda & BU & BR & S \end{bmatrix} \succ 0, \quad (24)$$

where $\Omega = CS = C_o S + D_u J$ and $\Lambda = AS = FS + GJ$. Similarly, condition (19) in Theorem 2 can be written as

$$\begin{bmatrix} S & -\Omega^\top & \Lambda^\top \\ -\Omega & 2(H+I)U & UB^\top \\ \Lambda & BU & S \end{bmatrix} \succ 0, \quad (25a)$$

$$\begin{bmatrix} S & \Omega_{i,:}^\top \\ \Omega_{i,:} & \tilde{y}_i^2(h_i) \end{bmatrix} \succeq 0, \quad \forall i \in \{1, \dots, \nu\}, \quad (25b)$$

Besides stability, the state-feedback gain K should be chosen to impose desired closed-loop performances. Consider the linearization of (2) around zero equilibrium

$$x^+ = (F + GK)x + d. \quad (26)$$

The term d can be regarded as an exogenous disturbance or, recalling that $d = Bq(y)$, it can account for neglected nonlinearities, whose effect can be attenuated by means of suitable minimization-based design approaches. As later discussed, different control strategies can be combined with the conditions described in Section 3 to enforce performance and stability properties to the closed-loop system via matrix inequalities. Here, the \mathcal{H}_2 -based approach is leveraged, given its desirable noise rejection properties. In particular, \mathcal{H}_2 control minimizes the effect of the exogenous disturbance d , representing the nonlinearity excess (see (2)), on the performance output

$$z = \tilde{Q}x + \tilde{R}u = (\tilde{Q} + \tilde{R}K)x, \quad (27)$$

where \tilde{Q}, \tilde{R} are tunable output matrices. K is designed to minimize the \mathcal{H}_2 norm of the transfer function from d to z in the region of linear behavior, namely in the tail of the nonlinear closed-loop responses. The next proposition, whose proof is reported in La Bella et al. (2024), allows to cast the design as a quasi-convex optimization problem.

Proposition 2. The \mathcal{H}_2 control problem applied to (26) can be recast as the solution to the following (quasi-convex) generalized eigenvalue problem

$$\min_{\substack{J \in \mathbb{R}^{m \times n}, \Gamma = \Gamma^\top \in \mathbb{R}^{n \times n} \\ S \in \mathbb{S}_{>0}^n, \delta \in \mathbb{R}_{>0}, \eta \in \mathbb{R}_{>0}}} \delta$$

subject to

$$\begin{bmatrix} S & SF^\top + J^\top G^\top & S\tilde{Q}^\top + J^\top \tilde{R}^\top \\ FS + GJ & S & 0 \\ \tilde{Q}S + \tilde{R}J & 0 & \eta I \end{bmatrix} \succ 0, \quad (28a)$$

$$\begin{bmatrix} \Gamma & \eta I \\ \eta I & S \end{bmatrix} \succeq 0, \quad (28b)$$

$$\text{trace}(\Gamma) < \eta \delta, \quad (28c)$$

where the state-feedback gain is selected as $K = JS^{-1}$.

Note that, if $\eta = 1$, (28) becomes an LMI-based problem similar to van Waarde et al. (2020). This would however restrict the degrees of freedom of the problem, limiting the size of the estimate of the basin of attraction $\mathcal{E}(S)$. Due to Proposition 2, in order to guarantee desired performances, the local stability of the origin relying on Theorem 1, and the maximization of the basin of attraction estimate, the following problem can be formulated

$$\min_{\substack{S, \Gamma \in \mathbb{S}_{>0}^n, L \in \mathbb{R}^{\nu \times n}, U, R \in \mathbb{D}_{>0}^\nu \\ J \in \mathbb{R}^{m \times n}, \delta \in \mathbb{R}_{>0}, \eta \in \mathbb{R}_{>0}}} w \delta - \gamma \quad (29)$$

subject to (9b), (24), (28), and $f_{\text{LMI}}(S) \geq \gamma I$,

where $w \in \mathbb{R}_{\geq 0}$ is a design parameter, which can be tuned based on the prioritization of the \mathcal{H}_2 norm minimization or the basin of attraction maximization, and $f_{\text{LMI}}(S)$ a function selected by the designer as discussed in Section 3.2. On the other hand, we can address the control design problem by resorting on Theorem 2. Given the nonlinearity (25), a recursive procedure as the one sketched in Algorithm 1 must be adopted. In particular, it is enough to solve, in step 5, the following alternative design problem

$$\min_{\substack{J \in \mathbb{R}^{m \times n}, S, \Gamma \in \mathbb{S}_{>0}^n \\ U \in \mathbb{D}_{>0}^n, \delta \in \mathbb{R}_{>0}, \eta \in \mathbb{R}_{>0}}} w \delta - \gamma \quad (30)$$

subject to (25), (28), and $f_{\text{LMI}}(S) \geq \gamma I$,

with $w \in \mathbb{R}_{>0}$ being a design parameter, and $f_{\text{LMI}}(S)$ a suitably chosen function for the maximization of the estimate of the basin of attraction. The next main result is a straightforward consequence of Theorems 1 and 2.

Theorem 3. Any solution to (29) or (30) is such that the origin is exponentially stable for the closed-loop system (2) with the selection $K = JS^{-1}$, with a basin of attraction containing the set $\mathcal{E}(S)$.

Remark 4. Denoting by J^*, S^* the optimal solution either of (29) or (30), it is worth noting that the resulting estimate of the basin of attraction, i.e., $\mathcal{E}(S^*) = \{x : x^\top S^{*-1} x \leq 1\}$, may be not the largest one, as this is not maximized in the \mathcal{H}_2 control problem. However, a larger basin of attraction can be computed by selecting the optimal state-feedback gain as $K = J^* S^{*-1}$, and solving either (16) or Algorithm 1 following the analysis results in Section 3.

Remark 5. Other control design strategies can be combined with the proposed regional stability conditions, leading to LMI control design problems, such as \mathcal{H}_∞ , using ideas from (van Waarde et al., 2020), or Virtual Reference Feedback Tuning (VRFT), using ideas from (D'Amico et al., 2023a).

5. Numerical results

The proposed approach is tested on a realistic benchmark taken from the literature, representing a pH neutralization process; for details see Bugliari Armenio et al. (2019). The pH process is a nonlinear SISO dynamical system where the input is the alkaline flowrate and the output is the pH concentration of a chemical solution. This system is simulated on laptop with an Intel i7-6500U processor, whereas the optimization problems are solved using Matlab with MOSEK as a solver.

First of all, the system model is exploited to generate input-output data through different experiments. In particular, 30000 samples of data are collected with a sampling time $T_s = 25$ s, using a multilevel pseudo-random signal (MPRS) for the input (i.e., the alkaline flowrate) with amplitude in the range [14.35, 16.35] mL/s. This dataset is exploited to derive a ESN model identifying the system dynamics, having structure

$$\begin{cases} x_s^+ &= \sigma_s(W_x x_s + W_u u + W_{xy} y_s), \\ y_s &= W_y x_s, \end{cases} \quad (31)$$

where $x_s \in \mathbb{R}^{n_s}$, $u \in \mathbb{R}$ and $y_s \in \mathbb{R}$ are the ESN state, input, and output vectors, respectively, the function $\sigma_s : \mathbb{R}^{n_s} \mapsto \mathbb{R}^{n_s}$ has components $\sigma_{s,i}(\cdot) = \tanh(\cdot)$ for all $i = 1, \dots, n_s$, whereas W_x , W_u , W_{xy} , and W_y are the weighting matrices. The latter are tuned exploiting standard training procedures for ESNs, see Bugliari Armenio et al. (2019) for more details, selecting 70% of data as training set and 30% as validation set. The model order has been

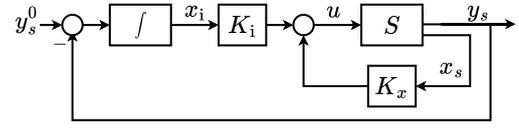


Fig. 4. State-feedback control scheme with integral action.

selected as $n_s = 3$, which achieves a FIT equal to 73% on the validation set. Note that the control design problem (29) displays numerical problems for ESN models with $n_s > 3$. On the other hand, the control design problem (30) was always able to find a feasible solution for any of the tested number of ESN states with $n_s \in \{3, 50\}$. Note that advanced techniques could be exploited for solving large-scale matrix problems, such as those based on matrix sparsity properties (Zhang & Lavaei, 2018) or on second order cone programming (Ahmadi et al., 2017).

The objective of the control design procedure is to tune the gains K_x and K_i of the state-feedback and integrator which are, respectively, components of the control input u , as shown in Fig. 4. The integrator has model $x_i^+ = x_i + e = x_i + y_s^0 - y_s$, where y_s^0 is the output reference to be tracked and $e = y_s^0 - y_s$ is the tracking error. Defining the closed-loop state vector as $x = [x_s^\top, x_i]^\top$ and the control input with a state-feedback control law, i.e., $u = [K_x K_i] x = Kx$, the overall closed-loop system model can be written in the form

$$\begin{cases} x^+ &= (F + GK)x + B q(y) + B_r y_s^0, \\ y &= Cx, \end{cases} \quad (32)$$

where $F = \begin{bmatrix} W_x + W_{xy} W_y & 0 \\ -W_y & 1 \end{bmatrix}$, $G = [W_u^\top \ 0]^\top$, $B = [-I \ 0]^\top$, $B_r = [0 \ 1]^\top$, $C = [W_x + W_{xy} W_y + W_u K_x \ W_u K_i]$, and $q(y) = y - \sigma_s(y)$. It is apparent that (32) has the same structure as (2), apart from the presence of the reference input y_s^0 , which has no influence on the stability and control design conditions described in the previous sections. Note that the ESN state x_s in (31) is not measurable in practice. Hence, a state observer is employed and designed as discussed in D'Amico et al. (2023a), enjoying δ ISS, and thus, vanishing estimation error (see also Remark 3). In the setting above, the control design solutions (29) and (30) are applied, so as to compute an optimal gain K ensuring \mathcal{H}_2 performance and the stability of the origin for the nonlinear closed-loop system (32). As for Theorem 3, the \mathcal{H}_2 performance output z is selected as $z_k = [\sqrt{\tilde{q}_y} y^\top \ \sqrt{\tilde{q}_i} x_i^\top \ \sqrt{\tilde{r}_u} u^\top]^\top$. In particular, the \mathcal{H}_2 performance output is selected so as to minimize the output tracking error and the integrator state through the weights \tilde{q}_y and \tilde{q}_i , as well as the control input effort through the weight \tilde{r}_u .

Thus, the proposed \mathcal{H}_2 control problem is solved by setting $\tilde{q}_y = \tilde{q}_i = 0.1$, $\tilde{r}_u = 0.05$, and $w = 1$. Note that \mathcal{H}_2 control problem applied in combination with the global stability condition (4) is not able to find a feasible solution. This is due to the presence of the system integrator and the bounded characteristic of the output in (31), as discussed in D'Amico et al. (2024). Thus, regional conditions are necessary, which can also define the maximum basin of attraction where the closed-loop stability property holds, by relying on the control problems (29) and (30), which are solved by setting $f_{\text{LMI}}(S) = S$. As apparent from the cost functions used in (29) and (30), the proposed design problems are multi-objective. In this section, to properly analyze the trade-off between the achieved performances, evaluated in terms of the \mathcal{H}_2 norm, and the achieved dimension of the basin of attraction, evaluated in terms of the radius of $\mathcal{E}(S)$, the control design problems (29) and (30) are solved by fixing the value of δ , i.e., setting $\delta = \bar{\delta}$ with $\bar{\delta} \in [5, 20]$. The achieved results are depicted in Fig. 5, where the trade-off between the performances that can be requested to

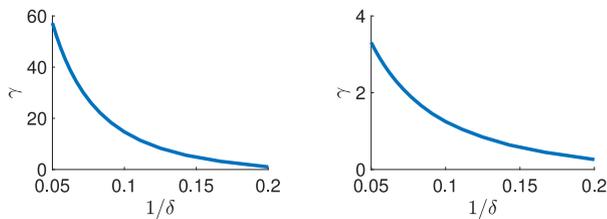


Fig. 5. Radius of the ellipsoid $\varepsilon(S)$, i.e., γ , for varying \mathcal{H}_2 norm $\bar{\delta} \in [5, 20]$. Left plot: solution obtained from (29), right plot: solution obtained from (30).

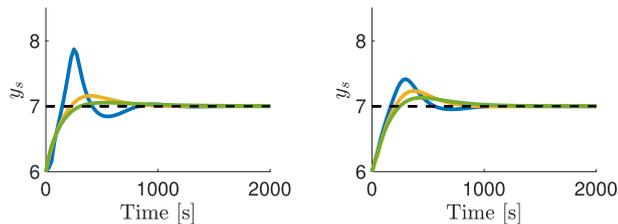


Fig. 6. Left plot: Output response with control law computed from (29), right plot: computed from (30) for $\bar{\delta} = 5$ (blue), $\bar{\delta} = 10$ (yellow), $\bar{\delta} = 20$ (green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the control system and the achieved basin of attraction becomes apparent. In fact, the larger $1/\delta$, i.e., the smaller the \mathcal{H}_2 norm, the smaller the radius of the estimate $\varepsilon(S)$ of the basin of attraction. Also, notably, the control solution (29) attains a larger estimate of the basin of attraction, with comparable \mathcal{H}_2 norm. Finally, the designed closed-loop controllers are tested on the real pH process model, considering a reference to be tracked at $y_k^0 = 7$; the results are reported in Fig. 6. The latter shows that (for the controllers obtained with (29) and (30)) the lower the \mathcal{H}_2 norm the smoother the output response, with reduced overshoots in the output transient.

6. Conclusions

In this paper novel global and regional stability analysis conditions based on LMIs have been proposed for a general class of RNNs. In particular, two alternative conditions are studied; the first one is obtained by defining an auxiliary function, while the second one is obtained by a sector narrowing technique. We have shown how to leverage these conditions for state-feedback control design by formulating a suitable optimization problem enforcing \mathcal{H}_2 norm minimization. The theoretical properties of the methods presented here have been thoroughly characterized. Finally, numerical simulations showed the advantages and limitations of the two methods. In particular, among the two regional stability conditions, the former one (based on the auxiliary function) provides a larger estimate of the basin of attraction as compared to the latter one (based on sector narrowing) when both are feasible, but numerical tests show that the former condition may lead to numerical issues when high-dimensional systems are employed. On the other hand, the condition based on sector narrowing, despite returning a smaller estimate of the basin of attraction, manages to find a control design solution with stability guarantees also when large-order systems are considered. This trade-off will be studied in future works. Future research will also include the extension of the obtained conditions to more complex control systems, possibly including nonlinear regulators belonging to more advanced RNN architectures (e.g., LSTM or GRU).

Acknowledgments

The research activity has been carried out within the project “SafeTy Aware Reinforcement Learning for robotic inspection” (STARLIT) funded by the Italian MUR, CUP n. E53D23001130006, number 2022ZE9J9J and by the ANR via grant OLYMPIA, number ANR-23-CE48-0006.

References

- Ahmadi, A. A., Dash, S., & Hall, G. (2017). Optimization over structured subsets of positive semidefinite matrices via column generation. *Discrete Optimization*, 24, 129–151.
- Barabanov, N. E., & Prokhorov, D. V. (2002). Stability analysis of discrete-time recurrent neural networks. *IEEE Transactions on Neural Networks*, 13(2), 292–303.
- Bayer, F., Bürger, M., & Allgöwer, F. (2013). Discrete-time incremental ISS: A framework for robust NMPC. In *2013 European control conference* (pp. 2068–2073). IEEE.
- Bonassi, F., Farina, M., & Scattolini, R. (2021). On the stability properties of gated recurrent units neural networks. *Systems & Control Letters*, 157, Article 105049.
- Bonassi, F., Farina, M., Xie, J., & Scattolini, R. (2022). On recurrent neural networks for learning-based control: recent results and ideas for future developments. *Journal of Process Control*, 114, 92–104.
- Bonassi, F., La Bella, A., Panzani, G., Farina, M., & Scattolini, R. (2023). Deep long-short term memory networks: Stability properties and experimental validation. In *2023 European control conference* (pp. 1–6).
- Boyd, S., El Ghaoui, L., Feron, E., & Balakrishnan, V. (1994). *Linear matrix inequalities in system and control theory*. SIAM.
- Buehner, M., & Young, P. (2006). A tighter bound for the echo state property. *IEEE Transactions on Neural Networks*, 17(3), 820–824.
- Bugliari Armenio, L., Terzi, E., Farina, M., & Scattolini, R. (2019). Model predictive control design for dynamical systems learned by echo state networks. *IEEE Control Systems Letters*, 3(4), 1044–1049.
- Chlipala, A. (2022). *Certified programming with dependent types: a pragmatic introduction to the Coq proof assistant*. MIT Press.
- Da Silva, J. G., & Tarbouriech, S. (2001). Local stabilization of discrete-time linear systems with saturating controls: an LMI-based approach. *IEEE Transactions on Automatic Control*, 46(1), 119–125.
- D'Amico, W., La Bella, A., Dercole, F., & Farina, M. (2023a). Data-based control design for nonlinear systems with recurrent neural network-based controllers. *IFAC-PapersOnLine*, 56(2), 6235–6240.
- D'Amico, W., La Bella, A., & Farina, M. (2024). An incremental input-to-state stability condition for a class of recurrent neural networks. *IEEE Transactions on Automatic Control*, 69(4), 2221–2236. <http://dx.doi.org/10.1109/TAC.2023.3327937>.
- de Souza, C., Girard, A., & Tarbouriech, S. (2023). Event-triggered neural network control using quadratic constraints for perturbed systems. *Automatica*, 157, Article 111237.
- de Souza, C., Tarbouriech, S., & Girard, A. (2023). Event-triggered neural network control for LTI systems. *IEEE Control Systems Letters*, 7, 1381–1386.
- Goebel, R., Sanfelice, R., & Teel, A. (2012). *Hybrid dynamical systems: Modeling, stability, and robustness*. London: Princeton University Press.
- Jiang, Z.-P., & Wang, Y. (2001). Input-to-state stability for discrete-time nonlinear systems. *Automatica*, 37(6), 857–869.
- La Bella, A. (2024). *Sector bound verification via COQ*. GitHub, https://github.com/alessiolabella/COQproof_RegionalStability.
- La Bella, A., Farina, M., D'Amico, W., & Zaccarian, L. (2024). Regional stability conditions for recurrent neural network-based control systems. URL <http://arxiv.org/abs/2409.15792>.
- Lanzetti, N., Lian, Y. Z., Cortinovis, A., Dominguez, L., Mercangöz, M., & Jones, C. (2019). Recurrent neural network based MPC for process industries. In *2019 18th European control conference* (pp. 1005–1010). IEEE.
- Liu, P., Wang, J., & Zeng, Z. (2021). An overview of the stability analysis of recurrent neural networks with multiple equilibria. *IEEE Transactions on Neural Networks and Learning Systems*.
- Massimetti, M., Zaccarian, L., Hu, T., & R. Teel, A. (2009). Linear discrete-time global and regional anti-windup: an LMI approach. *International Journal of Control*, 82(12), 2179–2192.
- Revay, M., Wang, R., & Manchester, I. R. (2020). A convex parameterization of robust recurrent neural networks. *IEEE Control Systems Letters*, 5(4), 1363–1368.

- Sontag, E. D. (1992). Neural nets as systems models and controllers. In *Proc. seventh yale workshop on adaptive and learning systems* (pp. 73–79).
- Stipanović, D. M., Kapetina, M. N., Rapaić, M. R., & Murmann, B. (2021). Stability of gated recurrent unit neural networks: Convex combination formulation approach. *Journal of Optimization Theory and Applications*, 188(1), 291–306.
- Vance, J., & Jagannathan, S. (2008). Discrete-time neural network output feedback control of nonlinear discrete-time systems in non-strict form. *Automatica*, 44(4), 1020–1027.
- van Waarde, H. J., Camlibel, M. K., & Mesbahi, M. (2020). From noisy data to feedback controllers: Nonconservative design via a matrix S -lemma. *IEEE Transactions on Automatic Control*, 67(1), 162–175.
- Yin, H., Seiler, P., & Arcaç, M. (2021). Stability analysis using quadratic constraints for systems with neural network controllers. *IEEE Transactions on Automatic Control*, 67(4), 1980–1987.
- Zhang, R. Y., & Lavaei, J. (2018). Efficient algorithm for large-and-sparse LMI feasibility problems. In *2018 IEEE conference on decision and control* (pp. 6868–6875). IEEE.



Alessio La Bella received the B.Sc. and M.Sc. in Automation Engineering cum laude from Politecnico di Milano in 2013 and 2015, respectively, and the M.Sc. in Mechatronics Engineering cum laude at Politecnico di Torino in 2016. He received the Ph.D. Degree cum laude in Information Technology at Politecnico di Milano in 2020. He was Visiting Researcher at the Automatic Control Lab of the École Polytechnique Fédérale de Lausanne, Switzerland, in 2018, and at the Center for Systems and Control of the Technische Universiteit Delft, The Netherlands, in 2024. From 2020 to 2022, he worked as Research Engineer at Ricerca sul Sistema Energetico - RSE SpA, designing and deploying advanced predictive controllers for energy plants in collaboration with industrial companies and utilities. In 2022, he joined as Assistant Professor the Department of Electronics, Information and Bioengineering of Politecnico di Milano. He is Associate Editor of the International Journal of Adaptive Control and Signal Processing and of EUCA Conference Editorial Board. His research interests concern the theory and design of predictive, multi-agent and learning-based control systems, with particular emphasis on practical challenges arising from the upcoming energy transition. He was recipient of the Dimitris N. Chorafas Prize in 2020.



Marcello Farina received the M.Sc. degree in Electronic Engineering in 2003 and the Ph.D. degree in Information Engineering in 2007, both from Politecnico di Milano. In 2005 he was visiting student at the Institute for Systems Theory and Automatic Control, Stuttgart, Germany. Presently, he is Associate Professor at Dipartimento di Elettronica e Informazione, Politecnico di Milano. His theoretical research interests include model predictive control, distributed and decentralized state estimation and control, data-based methods for direct and indirect design of control systems, and the use of recurrent neural network models in controller design. His applicative research includes the development of assistive devices.



William D'Amico received the M.Sc. degree cum laude in Automation and Control Engineering in 2019 and the Ph.D. degree cum laude in Information Technology in 2024, both from Politecnico di Milano, Italy. In 2022 and 2023, he was a visiting Ph.D. student with the Engineering and Technology Institute of the University of Groningen, The Netherlands. In 2024, he received the annual award for Ph.D. theses in Systems and Control Engineering from SIDRA, the Italian Control Systems Society. His research interests include data-driven control, recurrent neural network models, and deep learning.



Luca Zaccarian received the Laurea and the Ph.D. degrees from the University of Roma Tor Vergata (Italy) in 1995 and 2000, respectively. He has been Assistant Professor in control engineering at the University of Roma, Tor Vergata (Italy), from 2000 to 2006 and then Associate Professor. Since 2011 he is Directeur de Recherche at the LAAS-CNRS, Toulouse (France) and since 2013 he holds a part-time professor position at the University of Trento, Italy. Luca Zaccarian's main research interests include analysis and design of nonlinear and hybrid control systems, modeling and control of mechatronic systems. He is currently senior editor for nonlinear systems and control for the IFAC journal Automatica, a member of the EUCA-CEB and an associate editor for the European Journal of Control. He was a recipient of the 2001 O. Hugo Schuck Best Paper Award given by the American Automatic Control Council. He is a fellow of the IEEE, class of 2016.