LEARNING DYNAMICAL SYSTEMS WITH REGIONAL STABILITY CERTIFICATES

Anonymous authors
Paper under double-blind review

ABSTRACT

Regionally stable systems produce bounded outputs only on part of their state and input space, outside it, their outputs grow unbounded even for bounded inputs. Unconstrained models can reproduce such behavior but provide no guarantees, whereas globally stable models provide guarantees but cannot predict unbounded outputs. We propose a dynamic recurrent neural network in Lur'e form with deadzone nonlinearity. Exploiting the deadzone, we derive sufficient conditions on the learnable parameters that yield a regional stability certificate, which consists of an invariant state set, a state-dependent input map, and an output set that contains the model predictions whenever the initial state and inputs are admissible. During training, we enforce these conditions with logarithmic barrier functions and enlarge the admissible input map with a semidefinite program. On two synthetic systems, the learned model recovers the regional stability properties of the data-generating system and reproduces its unbounded response. On datasets from a nonlinear system identification benchmark that do not contain unbounded outputs, they match the accuracy of globally stable models, while unconstrained models are more accurate but uncertified.

1 INTRODUCTION

A ship rolling in open water returns upright as long as waves and wind stay moderate, but capsizes once a strong wind pushes it past a critical angle (Wawrzynski, 2021). We call such dynamical systems, with a bounded and unbounded regime, *regionally stable*. In this work, we learn models of regionally stable systems from input-output data, along with a certificate stating for which initial states and inputs the model is guaranteed to remain bounded.

Models of dynamic systems are used to simulate, design, and control them and require accurate predictions (Schoukens & Ljung, 2019). For regionally stable systems, these models are only useful if they capture the bounded and unbounded regimes. A model that predicts bounded responses everywhere suggests a safety margin that does not exist (Revay et al., 2020; Gu et al., 2022; Revay et al., 2021). A model that predicts unbounded responses inside the stable region might raise false alarms (Beintema et al., 2023b; Mohajerin & Waslander, 2019). A model of a regionally stable system should therefore provide a certificate for a set of inputs and initial states for which the predictions are ensured to remain bounded. In a downstream task, such a model certificate can be used to design a controller and ensure closed-loop stability (Bonassi et al., 2024).

To achieve high prediction accuracy and ensure stability, we propose a regionally stable dynamic recurrent neural network (RNN) with deadzone activation functions. We derive conditions on the model's learnable parameters that ensure the predicted output is bounded whenever the input stems from a state-dependent input map and the initial condition is from a compact set. In addition to the learned model parameters, we characterize the input map, state, and output sets.

We evaluate our model on two synthetic, regionally stable systems and five system-identification benchmarks (Schoukens et al., 2026). On synthetic systems, our model recovers regional behavior, outperforms models that are stable everywhere, and matches the accuracy of unconstrained models. On benchmarks whose data contain no unbounded outputs, it matches the accuracy of models that are stable everywhere while providing a certificate. Unconstrained models are more accurate but uncertified.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Our contributions are:

- a novel definition of a regional stability certificate that consists of an invariant state set, an admissible (state-dependent) input map, and an invariant output set (Section 3),
- sufficient conditions on the model parameters, in the form of linear matrix inequality constraints, that guarantee the learned model admits such a regional certificate (Section 4),
- a training algorithm that minimizes the prediction error, while ensuring the parametric constraints are satisfied, and the admissible input map is maximized (Section 5).

We position our work in the current literature in Section 2. The numerical experiments are presented in Section 6, and we provide concluding remarks in Section 7.

NOTATION

We denote by $\mathbb{R}_+$ ($\mathbb{N}_+$) the nonnegative real (natural) numbers. For $x \in \mathbb{R}^n$, we define $|x|^2 := x^{\mathrm{T}}x$, and x^i denotes the i-th element of x. For a sequence $x = (x_k)_{k\in\mathbb{N}_0}$, with $x_k \in \mathbb{R}^n$, we define $\|x\|_\infty := \sup_k |x_k|$ and the signal space $\ell_\infty^n := \{x \mid \|x\|_\infty < \infty\}$. We denote by $\mathbb{S}_+^n$ the symmetric positive definite $n \times n$ matrices and by $\mathbb{D}_+^m$ the diagonal positive definite $m \times m$ matrices. The notation $A \succ 0$ ($A \prec 0$) means that the matrix A is positive definite (negative definite). We use $(\star)$ for matrices or vectors that can be obtained from symmetry. For $Q \in \mathbb{S}_+^n$, we define the ellipsoid $\mathcal{E}(Q) := \{x \in \mathbb{R}^n : x^{\mathrm{T}}Qx \leq 1\}$, and for $H \in \mathbb{R}^{q\times n}$ with rows h^i the polyhedral set $\mathcal{L}(H) := \{x \in \mathbb{R}^n : \|Hx\|_\infty \leq 1\}$. A function $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ is of class $\mathcal{K}$ if it is continuous, strictly increasing, and $\gamma(0) = 0$, and of class $\mathcal{K}_\infty$ if, in addition, $\lim_{s\to\infty} \gamma(s) = \infty$. A function $\beta : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is of class $\mathcal{KL}$ if $\beta(\cdot, k)$ is of class $\mathcal{K}$ for each $k \geq 0$, and $\beta(s, \cdot)$ is nonincreasing with $\lim_{k\to\infty} \beta(s, k) = 0$ for each $s \geq 0$ for each $s \geq 0$ (Jiang & Wang, 2001).

2 RELATED WORK

System identification Learning models of dynamical systems from input-output measurements is known as system identification (Zadeh, 1962; Ljung, 1998). Since linear models (Overschee & Moor, 1996) are limited in expressiveness, nonlinear model classes have been developed (Schoukens & Ljung, 2019), neural networks are a subclass and have achieved high prediction accuracy (Ljung, 1998; Pillonetto et al., 2025; Beintema et al., 2023a; Hu et al., 2026). RNNs (Goodfellow et al., 2016) capture long-term dependencies between inputs and outputs, which makes them well-suited for learning dynamical systems (Mohajerin & Waslander, 2019; Baier et al., 2023). We build on them to learn regionally stable models.

Stable system identification Enforcing stability of nonlinear models is difficult, since their dynamics can exhibit multiple equilibria, limit cycles, and unbounded solutions (Khalil, 2002; Sastry, 2013). Suykens et al. (1995) showed that RNNs can be written as uncertain linear state-space models, which makes tools from robust control applicable (Scherer et al., 1997). For feed-forward networks, this yields linear matrix inequalities (LMI) conditions for Lipschitz bounds (Fazlyab et al., 2019; Pauli et al., 2021b). For RNNs in system identification, (Revay et al., 2020; 2021; Gu et al., 2022) guarantee contraction, i.e., trajectories from different initial states converge to each other under the same input, and Bonassi et al. (2022; 2024) derive input-to-state stable (ISS) for a family of recurrent networks, e.g., LSTMs and GRUs, which are more conservative than LMI-based constraints (Frank et al., 2026). Both guarantees hold globally and therefore rule out unbounded predictions. Regional guarantees have been derived for closed-loop systems with neural network controllers, where local sector conditions yield an inner approximation of the region of attraction (Yin et al., 2021), and for kernel-based models via regional convergence (Pavlov et al., 2004; Shakib et al., 2023), verified on a grid of LMIs. In contrast, we learn an RNN from input-output data and certify it, together with a state-dependent map of admissible inputs, by a fixed number of LMIs.

Generalized sector conditions Our certificate relies on generalized sector conditions, which originate in anti-windup design for feedback loops with saturation (Da Silva & Tarbouriech, 2005; Tarbouriech et al., 2006; Massimetti et al., 2009; Tarbouriech et al., 2011). La Bella et al. (2025) use them to analyze autonomous closed loops in which a neural network acts as the controller. Applying

them to system identification requires handling exogenous model inputs, which we address with a state-dependent input map.

3 PROBLEM FORMULATION

3.1 DATA-GENERATING SYSTEM, DATA AND MODEL CLASS

A data-generating system is a nonlinear system Σ that maps an input sequence u to an output sequence y.

We observe the data generating system Σ through finite-length input-output trajectories. The dataset is defined as follows:

$$\mathcal{D} = \left\{ \left(y_k^{(i)}, u_k^{(i)} \right)_{k=0}^{N_i - 1} \right\}_{i=1}^{M}, \tag{1}$$

constituting M trajectories, each of length N_i. In (1), we denote the i-th pair of input-output sequences by $(y, u)^{(i)}$.

We aim to identify the data-generating system Σ from $\mathcal{D}$ with a nonlinear discrete-time dynamic state-space model in Lur'e form

$$G := \left\{ \begin{pmatrix} x_{k+1} \\ \hat{y}_k \\ v_k \end{pmatrix} = \begin{pmatrix} A & B & B_2 \\ C & 0 & 0 \\ C_2 & D_{21} & 0 \end{pmatrix} \begin{pmatrix} x_k \\ u_k \\ w_k \end{pmatrix} \right. \qquad \text{with} \qquad w_k = \begin{pmatrix} \psi^1(v_k^1) \\ \vdots \\ \psi^q(v_k^q) \end{pmatrix} = \Delta(v_k), \tag{2}$$

and internal state $x_k \in \mathbb{R}^n$, input sequence u with $u_k \in \mathbb{R}^m$, and output $\hat{y}$ with $\hat{y}_k \in \mathbb{R}^p$. The nonlinear block Δ consists of the repeated nonlinear activation function $\psi : \mathbb{R} \to \mathbb{R}$ with $\psi^i = \psi$ for all $i = 1, \ldots, q$ that is applied elementwise. We call the model (2) a dynamic RNN (Goodfellow et al., 2016; Suykens et al., 1995). The decomposition of model (2) isolates the nonlinearity in Δ from the linear time-invariant block G, which allows us to use tools from robust control Scherer & Weiland (2000) to study its stability properties.

The learnable parameters are all in the linear block G of model (2) and collected as follows:

$$\theta = \{A,\ B,\ B_2,\ C,\ C_2,\ D_{21}\}. \tag{3}$$

The dynamic RNN model structure $\mathcal{M}$ is a set of dynamic RNNs where the parameters are from a set $\Theta \subseteq \mathbb{R}^{n_\theta}$. Each such nonlinear model $\mathcal{S}(\theta) \in \mathcal{M}$ maps an initial state x_0 and an input sequence u to a predicted output sequence $\hat{y}$. The overall aim is to find a specific set of parameters θ of model (2) that approximate the data-generating system Σ accurately.

3.2 MODEL PROPERTIES

In addition to a predictive model, we aim to recover the stability properties of the data-generating system Σ. Our proposed regional stability certificate is based on a regional form of input-to-state stability. Thus, we recall the notion of global input-to-state stability for models of the form (2).

Definition 1 (Input-to-state stability (Jiang & Wang, 2001))**.** *Model* (2) *is* input-to-state stable *if there exist $\beta \in \mathcal{KL}$ and $\gamma \in \mathcal{K}_\infty$ such that for every initial condition $x_0 \in \mathbb{R}^n$ and every input sequence $u \in \ell_\infty^m$, the state trajectory satisfies*

$$|x_k| \leq \beta(|x_0|, k) + \gamma(|u|_\infty), \qquad \textit{for every} \qquad k \geq 0. \tag{4}$$

A model that is input-to-state stable (ISS) produces a bounded state trajectory for any bounded input. As evident from (4), the bound depends on the initial condition whose effect vanishes through the class $\mathcal{KL}$-function β, and the considered input sequence through the class $\mathcal{K}_\infty$-function γ.

Remark 1. *For zero input, i.e., $u_k = 0$ for all $k \geq 0$, a nonlinear model that is ISS has a globally asymptotically stable fixed point at $x = 0$.*

Input-to-state stability as defined in Definition 1 is a global property, the bound (4) is required to hold for every initial state and every bounded input. The dataset (1) that is available for learning,

however, consists of finitely many trajectories of finite length and therefore covers only a small part of the state and input range. Yet, outside the observed regime, bounded inputs may drive the system to arbitrarily large responses. Requiring global ISS of the model class (or other global stability requirements) would therefore rule out such behavior a priori and thus imposes a restriction that goes beyond what the data can support.

To capture models that can produce unbounded outputs for bounded inputs, we first introduce a state dependent input restriction as follows.

Definition 2 (Admissible input map)**.** *Let $\mathcal{X} \subset \mathbb{R}^n$ be a compact set with $0 \in \mathcal{X}$, and let $\mathcal{U} : \mathcal{X} \rightrightarrows \mathbb{R}^m$ be a uniformly bounded, set-valued map assigning each $x \in \mathcal{X}$ to a nonempty and compact set of admissible inputs $\mathcal{U}(x) \subseteq \mathbb{R}^m$.*

The intuition of the state-dependent input map is that a regionally stable system can be excited with a large input whenever the state is close to the origin. On the contrary, when the state is near the boundary of the set $\mathcal{X}$, only small inputs are allowed.

We connect the admissible input-map with the internal state by defining a set $\mathcal{X}$ that contains the internal state x_k for all $k \geq 0$ when the input is admissible.

Definition 3 (Forward invariance)**.** *Let $\mathcal{X} \subset \mathbb{R}^n$ be compact and $\mathcal{U}$ and admissible input map on $\mathcal{X}$. The set $\mathcal{X}$ is forward invariant for $\mathcal{S}(\theta)$ with respect to $\mathcal{U}$ if, starting from $\mathcal{X}$, the state trajectories of $\mathcal{S}(\theta)$ stay in $\mathcal{X}$ for all $x \in \mathcal{X}$ and all $u \in \mathcal{U}(x)$.*

The ISS definition 1 requires the input signal to be bounded, e.g., $\|u\|_\infty < \infty$, with the admissible input map $\mathcal{U}$ and the forward invariant state set $\mathcal{X}$ this is ensured (details can be found in the Appendix B.5).

For any dynamic RNN that is ISS, every class of bounded inputs admits a bounded forward invariant set, whose size depends on the input bound and on the set of initial states.

When a model $\mathcal{S}(\theta)$ is *forward invariant* and *input-to-state stable* for an *admissible input-map*, then we can ensure the predicted outputs also confine in a compact output set. The certificate is then defined by the corresponding sets.

Definition 4 (Regional certificate)**.** *Let $\mathcal{X} \subset \mathbb{R}^n$ be compact with $0 \in \mathcal{X}$, let $\mathcal{U} : \mathcal{X} \rightrightarrows \mathbb{R}^m$ be an admissible input map, and let $\mathcal{Y} \subset \mathbb{R}^p$ be compact. The triple $(\mathcal{X}, \mathcal{U}, \mathcal{Y})$ is a* regional certificate *for a dynamic RNN model of the form* (2) *if*

1. *the model is* forward invariant *on $\mathcal{X}$ with respect to $\mathcal{U}$,*
2. *there exists $\beta \in \mathcal{KL}$ and $\gamma \in \mathcal{K}_\infty$ such that, for every $x_0 \in \mathcal{X}$ and every input sequence satisfying $u_k \in \mathcal{U}(x_k)$ for all $k \geq 0$, the input-to-state stability bound* (4) *holds, and*
3. *$\hat{y}_k = Cx_k \in \mathcal{Y}$ for $x_k \in \mathcal{X}$ and $u_k \in \mathcal{U}(x_k)$ for all $k \geq 0$.*

We consider a model equipped with a *regional certificate* regionally stable.

One important distinction is that a regionally stable model may produce unbounded (states and) outputs for inputs outside the certified range, whereas globally stable model classes such as (Revay et al., 2020; Pauli et al., 2021b; Fazlyab et al., 2019; Shakib et al., 2022) exclude such behavior. In the Appendix A we show examples of the model properties introduced in this Section.

3.3 Admissible input size

In order to measure the size of the admissible input map we introduce the following size functional.

Definition 5 (Size functional)**.** *Let $\mathbf{U}$ denote the collection of admissible inputs maps with $0 \in \mathcal{U}(x)$ for all $x \in \mathcal{X}$. A size functional is a map $\sigma : \mathbf{U} \to \mathbb{R}_+$ that is monotone with respect to set inclusion, i.e., for any two admissible input maps $\mathcal{U}_1$ and $\mathcal{U}_2$ on a common state set $\mathcal{X}$*

$$\mathcal{U}_1(x) \subseteq \mathcal{U}_2(x) \qquad \textit{for all} \qquad x \in \mathcal{X} \implies \qquad \sigma(\mathcal{U}_1) \leq \sigma(\mathcal{U}_2). \tag{5}$$

In this work, we measure the size of an input map by largest admissible input for all $x_k \in \mathcal{X}$, defined as follows:

$$\sigma_u(\mathcal{U}) := \inf_{x \in \mathcal{X}} \sup\{\delta \geq 0 \mid \mathcal{B}_\delta \subseteq \mathcal{U}(x)\}, \qquad \mathcal{B}_\delta := \{u \in \mathbb{R}^m \mid \|u\| \leq \delta\}. \tag{6}$$

Lemma 1 (σ_u is a size functional)**.** *The map σ_u defined in* (6) *is a size functional for the input-map $\mathcal{U}$ according to Definition 5.*

The proof can be found in Appendix B.1. With the size functional we can now derive the largest admissible input attainable for $\mathcal{S}(\theta)$. Given parameter $\theta \in \Theta$, let $\bar{\sigma} : \Theta \to \mathbb{R}_+$ denote the largest admissible input attainable for $\mathcal{S}(\theta)$, with

$$\bar{\sigma}(\theta) \coloneqq \sup\{\sigma_u(\mathcal{U}) \in \mathbb{R} \mid (\mathcal{X}, \mathcal{U}, \mathcal{Y}) \text{ is a regional certificate for } \mathcal{S}(\theta)\}, \tag{7}$$

and $\bar{\sigma}(\theta) \coloneqq 0$ if $\mathcal{S}(\theta)$ if no input is certified.

3.4 Problem statement

Equipped with the notion of regional stability in Definition 4 and the size functional for the admissible input map (7), we can now state the learning problem.

Problem 1. *Given a dataset $\mathcal{D}$ as in* (1) *generated by an data-generating system Σ and a dynamic RNN model structure $\mathcal{M} := \{\mathcal{S}(\theta) \mid \theta \in \Theta\}$ with given parameter set Θ, find parameters $\theta \in \Theta$ corresponding to a model of the form* (2) *and a regional certificate $(\mathcal{X}, \mathcal{U}, \mathcal{Y})$ for $\mathcal{S}(\theta)$ such that*

- *the prediction error $\frac{1}{M}\sum_{i=1}^{M}\frac{1}{N_i}\sum_{k=0}^{N_i-1}\|y_k^{(i)} - \hat{y}_k^{(i)}\|^2$ is minimized, and*
- *the output set $\mathcal{Y}$ and the largest admissible input for all $x_k \in \mathcal{X}$ is maximized.*

Problem 1 aims to find model parameters θ that have minimal prediction error on the dataset (1) and maximize the output set and admissible input map in the sense that a maximal number of trajectories are covered by the regional certificate.

4 Regional Certificate for dynamic RNN

We introduce the deadzone activation function, which will be used as the activation function in the dynamic RNN model (2), and the generalized sector conditions it admits in Section 4.1. Then, we derive parametric constraints on the learnable dynamic RNN parameters under which the model is regionally stable according to Definition 4.

4.1 Generalized Sector Conditions

A scalar nonlinear function $\psi : \mathbb{R} \to \mathbb{R}$ in model (2) is said to satisfy *standard* sector conditions if

$$\begin{pmatrix} \psi(v) \\ v \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \psi(v) \\ v \end{pmatrix} \geq 0 \tag{8}$$

holds for all $v \in \mathbb{R}$, that is, if the graph ψ lies in the sector spanned by two lines through the origin with slope 0 and 1. Common activation function used in RNNs do not tightly cover the shape of ψ (see Appendix A.2 for details and plots). Stability results that rely on global sector conditions are therefore shown to be conservative (Pauli et al., 2021a; Scherer, 2023).

Models of the form (2) are universal function approximators (Suykens et al., 1995), and this property is retained when the $\tanh$ activation function used in RNNs is replaced by the saturation sat. A loop transformation (Khalil, 2002) then turns sat into the deadzone nonlinearity $\psi : \mathbb{R} \to \mathbb{R}$, defined by

$$\psi(v) \coloneqq \mathrm{dzn}(v) \coloneqq \begin{cases} 0 & \text{if} |v| \leq 1, \\ v - 1 & \text{if} v > 1, \\ v + 1 & \text{if} v < -1. \end{cases} \tag{9}$$

This step is exact since $\mathrm{sat}(v) = v - \mathrm{dzn}(v)$, the linear part of the substitution is absorbed into the matrices of the linear block G, leaving the input-output behavior of the model unchanged.

Using the dzn activation function (9) allows to generalize the sector conditions (8) by introducing a degree of freedom and link it to the model's internal state. To do so, we introduce the state-dependent quadratic form

$$\Gamma(v, x, \Lambda, H) \coloneqq \begin{pmatrix} \Delta(v) \\ v + Hx \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} -2\Lambda & \Lambda \\ \Lambda & 0 \end{pmatrix} \begin{pmatrix} \Delta(v) \\ v + Hx \end{pmatrix}, \tag{10}$$

which extends the *standard* sector condition by the following result.

Lemma 2 (Generalized sector conditions)**.** *Consider the nonlinear operator $\Delta : \mathbb{R}^q \to \mathbb{R}^q$ with activation function $\psi(z) = \mathrm{dzn}(z)$ applied elementwise, and matrices $\Lambda \in \mathbb{D}_+$ and $H \in \mathbb{R}^{q\times n}$. Then, for every $x \in \mathcal{L}(H)$ and $v \in \mathbb{R}^q$, we have*

$$\Gamma(v, x, \Lambda, H) \geq 0. \tag{11}$$

The proof is given in Appendix B.2. We refer to (11) as the *generalized* sector conditions. The set $\mathcal{L}(H) := \{x \in \mathbb{R}^n : \|Hx\|_\infty \leq 1\}$ defines the region in state-space where the generalized sector conditions hold. Conditions of this type are used in anti-windup controller design, see for example (Da Silva & Tarbouriech, 2005; Tarbouriech et al., 2006) and references therein.

4.2 Conditions for regional certificate

The following result is based on generalized sector conditions (Lemma 2) and provides sufficient conditions for a regional certificate. It characterizes the admissible input map $\mathcal{U}$ and the forward invariant state set $\mathcal{X}$

Theorem 1 (Regional input-to-state stability)**.** *Given a dynamic RNN model* (2) *with parameters $\theta \in \Theta$* (3) *and nonlinear operator $\Delta : \mathbb{R}^q \to \mathbb{R}^q$ with activation function $\psi(v) = \mathrm{dzn}(v)$ as in* (9) *applied elementwise, and consider a scalar $\alpha \in (0, 1)$. Suppose there exist matrices $P \in \mathbb{S}_+$, $M \in \mathbb{D}_+$, $L = \begin{pmatrix}(l^1)^{\mathrm{T}} & \cdots & (l)^q)^{\mathrm{T}}\end{pmatrix}^{\mathrm{T}} \in \mathbb{R}^{q\times n}$, and scalar $s \geq 0$ such that the LMI*

$$\underbrace{\left(\begin{array}{ccc|c} -\alpha^2 P & 0 & PC_2^{\mathrm{T}} + L^{\mathrm{T}} & PA^{\mathrm{T}} \\ \star & -I & D_{21}^{\mathrm{T}} & B^{\mathrm{T}} \\ \star & \star & -2M & MB_2^{\mathrm{T}} \\ \hline \star & \star & \star & -P \end{array}\right)}_{\mathbf{F}} \preceq 0, \tag{12a}$$

$$\underbrace{\begin{pmatrix} 1/s^2 & l^i \\ (l^i)^{\mathrm{T}} & P \end{pmatrix}}_{\mathbf{G}_i} \succeq 0, \forall i = 1, \ldots, q, \tag{12b}$$

are satisfied. Then, for any input from the admissible input map

$$u_k \in \mathcal{U}(x_k) \;\; \textit{with} \;\; \mathcal{U}(x_k) := \{u_k \in \mathbb{R}^m \mid (u_k)^{\mathrm{T}} u_k \leq s^2 - \alpha^2 x_k^{\mathrm{T}} X x_k\} \;\; \textit{for all} \;\; x_k \in \mathcal{X}, \tag{13}$$

the ellipsoid $\mathcal{E}(P^{-1}/s^2)$ is forward invariant and $\mathcal{E}(P^{-1}/s^2) \subset \mathcal{L}(H)$ with $H = LP^{-1}$ and $s > 0$. For any $x_0 \in \mathcal{E}(P^{-1}/s^2)$ and input u_k that satisfies (13)*, the following bound is satisfied for any $k \geq 0$:*

$$|x_k| \leq \min\Big\{\sqrt{\tfrac{\lambda_{\max}}{\lambda_{\min}}}\,\alpha^k |x_0| + \frac{\sup_{0\leq j<k} |u_j|}{\sqrt{\lambda_{\min}(1-\alpha^2)}}, \frac{s}{\sqrt{\lambda_{\min}}}\Big\}, \tag{14}$$

where λ_{min} and λ_{max} denote the minimum and maximum eigenvalues of the matrix P^{-1}. Moreover, if C has full row rank, the predicted output satisfies $\hat{y}_k \in \mathcal{Y} := \mathcal{E}\left((s^2 CPC^{\mathrm{T}})^{-1}\right)$ for all $k \geq 0$. Finally, the triplet $(\mathcal{X}, \mathcal{U}, \mathcal{Y})$ with $\mathcal{X} = \mathcal{E}(P^{-1}/s^2)$, $\mathcal{U}$ as in (13)*, and $\mathcal{Y} = \mathcal{E}\left((s^2 CPC^{\mathrm{T}})^{-1}\right)$ is a regional certificate in the sense of Definition 4.*

The proof is given in Appendix B.3. Theorem 1 introduces an additional set of learnable parameters that have no influence on the model prediction but shape the input map as well as the state and output set. We collect these parameters in a set $\kappa = (P, M, L, \hat{s}, \tau)$, with $\hat{s} = 1/s^2$ where the scalar $\alpha = \frac{1}{(1+e^{-\tau})} \in [0,\ 1)$ is obtained from the learnable τ. For fixed θ and α, the conditions (12) depend affinely on elements of κ, which allows to formulate feasibility problems as semidefinite programs (SDPs) that can be solved efficiently with off-the shelf SDP solvers. We will make use of that to restore feasibility during training and to find the largest admissible input map (in Section 5).

For the admissible input map we can define the size functional from Definition 5 based on the parameters s, P, and α, as illustrated in the following result.

Lemma 3 (Size functional for admissible input map of dynamic RNN)**.** *Let $\alpha \in (0, 1)$, $s > 0$, and $P \in \mathbb{S}_+^n$, and let $\mathcal{X} = \mathcal{E}(P^{-1}/s^2)$ with the admissible input map as in* (13)*, then*

$$\sigma_{\mathrm{u}}(\mathcal{U}) = s\sqrt{1-\alpha^2}. \tag{15}$$

The proof is given in Appendix B.4. The size functional for the admissible input map refers to the largest input that can be applied independent of the current state x_k as long as $x_k \in \mathcal{X}$.

When the output is one-dimensional, we define the output range as $\bar{y} := s\sqrt{CPC^{\mathrm{T}}}$, which means that for every admissible input $u \in \mathcal{U}$ and every initial state $x_0 \in \mathcal{X}$, the predicted model output is in the range $|y_k| \leq \bar{y}$ for all $k \geq 0$.

5 Training with regional stability certificate

Problem 1 asks for a model that is accurate and carries a large input-map. These two objectives are addressed by different mechanisms in our training algorithm. The prediction error is reduced by stochastic gradient descent (Section 5.1), while the certified range is enlarged by a semidefinite program that is solved between epochs with fixed θ and α (Section 5.2). Throughout training, the parametric constraints of Theorem 1 are kept satisfied, so that a certificate is available at every point of the optimization step. Details about initialization, ensuring feasibility, and the full algorithm can be found in the Appendix C.

5.1 Loss function

The loss function consists of a term penalizing prediction errors and a term promoting feasibility of the conditions (12). The conditions (12) are enforced through logarithmic barrier terms. For a real symmetric matrix $\mathbf{C}$, we define

$$\phi(\mathbf{C}) := \begin{cases} -\log\det(-\mathbf{C}) & \text{if } \mathbf{C} \prec 0, \\ \infty & \text{if } \mathbf{C} \not\prec 0, \end{cases} \tag{16}$$

so that ϕ grows without bound as the parameters approach the boundary of the feasible set. The loss on a trajectory $(y,u)^{(i)} \in \mathcal{D}$ is

$$\mathbf{L}^{(i)}(\theta,\kappa) = \frac{1}{N_i - w} \sum_{k=w}^{N_i-1} \left(\|\hat{y}_k^{(i)} - y_k^{(i)}\|^2 \right) + \nu\Big(\phi(\mathbf{F}) + \sum_{i=1}^{q} \phi(-\mathbf{G}_i)\Big), \tag{17}$$

where $\nu > 0$ weights the barrier against the fit. Both θ and κ are updated by gradient descent on (17).

5.2 Enlarging the certified range

The barrier term in (17) rewards distance from the constraint boundary rather than a large certified range. Therefore, the gradient descent step alone does not pursue the second objective of Problem 1. We address it by solving the optimization problem

$$\hat{s}^\star = \min_{P,M,L,\hat{s}} \hat{s} \quad \text{such that (12) holds} \tag{18}$$

with $\hat{s} := \frac{1}{s^2}$, which is a SDP for fixed θ and α. The solution to (18) attains $\bar{\sigma}(\theta)$ by Lemma 3 and we can derive $\bar{\sigma}(\theta) = \sqrt{\frac{1-\alpha^2}{\hat{s}^\star}}$.

Remark 2. *In* (18)*, we maximize the size of the admissible input map $\mathcal{U}$ attainable for a given θ. This is a proxy for enlarging the output range defined by $\mathcal{Y}$. Since θ is fixed, the input-output behavior of the model does not change, and a larger admissible input map certifies more input sequences and hence a larger range of outputs.*

6 Experiments

We compare three models of the form (2) that differ in the constraints imposed on the parameters during training: RegStable (ours), certified by the generalized sector conditions of Theorem 1; GlobStable, certified by the standard sector conditions used in globally stable parametrizations, which correspond to $H = 0$; and NotStable, trained without any stability constraints. The corresponding parameter sets are nested,

$$\Theta_{\text{GlobStable}} \subseteq \Theta_{\text{RegStable(ours)}} \subseteq \Theta_{\text{NotStable}}, \tag{19}$$

Table 1: Evaluation of regionally stable systems. Lowest NRMSE in **bold**, second lowest in ***bold+italic***; *# div/total*: diverging trajectories of the learned model / of the data-generating system (see Section 6.2 for details).

	One-dimensional system ($y_{\max} = 0.29$)				Softening Duffing oscillator ($y_{\max} = 0.934$)			
Model	NRMSE	# div/ total	$\bar{\sigma}(\theta)$	$\bar{y}$	NRMSE	# div/ total	$\bar{\sigma}(\theta)$	$\bar{y}$
LSTM	2.02	15/15	0	-	0.41	15/15	0	-
NOTSTABLE	**1.55**	15/15	0	-	**0.15**	4/15	0	-
GLOBSTABLE	2.22	0/15	∞	-	0.69	0/15	∞	-
REGSTABLE (ours)	***1.60***	15/15	3.13	0.888	***0.16***	7/15	$5.04 \cdot 10^{-5}$	0.358

so that the achievable simulation error is ordered in the opposite direction. The two baselines are the extreme cases of the regional certificate: as $\sigma_u(\mathcal{U}) \to \infty$ every input becomes admissible and the guarantee approaches that of a globally stable model (GLOBSTABLE), while NOTSTABLE carries no guarantee, the case $\bar{\sigma}(\theta) = 0$ in (7). Our model REGSTABLE (ours) interpolates between them.

In addition, we compare against an LSTM (LSTM), a standard recurrent architecture for system identification (Mohajerin & Waslander, 2019; Hu et al., 2026) which lies outside our model class (2).

6.1 DATASETS

Measurements from a real physical system rarely contain exponentially growing responses, since driving a system past its stable operating range risks damaging it. We therefore evaluate first on two systems whose governing equations are known and which are regionally but not globally stable.

Regionally stable systems: *One-dimensional system.* We use a one-dimensional system (see Appendix A for details) and generate 100 trajectories of length 500. The excitation is low-pass filtered white noise (Schoukens & Ljung, 2019), and the outputs are corrupted by additive noise at a signal-to-noise ratio of approximately 30 dB.

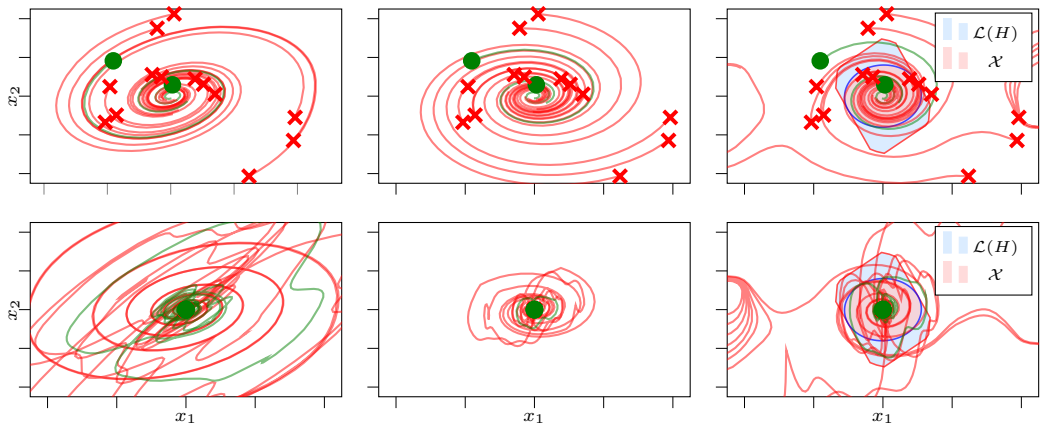
Softening Duffing oscillator. A two-dimensional system in which the cubic term carries the opposite sign to the linear term; these dynamics model, e.g., the roll motion of a ship in open water, and the state grows unbounded once the cubic nonlinearity dominates, corresponding to capsizing. We generate 100 trajectories of length 4000, discretized with step size of 0.05 again with low-pass filtered white-noise excitation and 30 dB output noise. For both regionally stable systems we generated 50 trajectories where the output grows exponentially.

Nonlinear benchmark A collection of nonlinear system identification problems (Schoukens et al., 2026), where the data is collected from real-world systems.

6.2 RESULTS

Accuracy on regionally stable systems. Table 1 reports the normalized root-mean-squared error (NRMSE) on held-out test trajectories, which contains both bounded and unbounded outputs. Consistent with the inclusion chain (19), REGSTABLE (ours) reaches an accuracy comparable to the unconstrained NOTSTABLE. Table 1 further shows that REGSTABLE (ours) is more accurate than GLOBSTABLE and LSTM.

Regional stability behavior. We test whether the learned models reproduce the stability behavior of the data-generating system by violating the input and initial state condition of the certificate separately: an initial state outside $\mathcal{X}$ with zero input, and an admissible initial state with an input outside $\mathcal{U}$. For the latter, we take the test inputs for which the data-generating system's output grows exponentially and pad them with zeros. A globally stable model returns to zero in both cases, whereas a regionally stable model may converge to a different fixed point or diverge (Remark 1). Figure 1 shows both tests for the softening Duffing oscillator in the phase plane: the simulation results depicted in the first row start from random initial states with admissible inputs $u \equiv 0$, whereas the simulation results in the second row start from admissible initial states but use inadmissible inputs.



(a) NOTSTABLE (b) GLOBSTABLE (c) REGSTABLE (ours)

Figure 1: Predictions of the learned models on the test set of the softening Duffing oscillator. Green indicates that the duffing oscillator is bounded, and red when the output grows unbounded. *Upper row*: random initial states ($u \equiv 0$) *Lower row*: inputs from the test set. For REGSTABLE (ours) in panel (c), the pink ellipse is the invariant set $\mathcal{X}$ and the blue region is $\mathcal{L}(H)$.

Table 2: NRMSE on datasets from the nonlinear benchmark Schoukens et al. (2026).

Model	Silverbox	EMPS	CED	F16	WH System
NOTSTABLE	**0.010**	**0.088**	0.33	0.54	0.162
GLOBSTABLE	0.023	***0.201***	0.37	**0.50**	***0.158***
LSTM	***0.013***	0.269	**0.27**	0.94	**0.005**
REGSTABLE (ours)	0.023	0.224	***0.31***	**0.50**	***0.158***

The column *# div/total* in Table 1 counts the trajectories that do not converge to zero, relative to the number of trajectories that diverge for the data-generating system. On both synthetic systems, REGSTABLE (ours) reproduces the behavior of the data-generating system, including its diverging trajectories, while GLOBSTABLE returns to zero even where the data-generating system diverges, as its global certificate implies. NOTSTABLE also reproduces the diverging trajectories but, like LSTM, provides no certificate, so it remains unknown for which initial states and inputs its predictions stay bounded. Details about the certified output $\bar{y}$ are in the Appendix D.

Nonlinear benchmarks. Table 2 reports the NRMSE on the benchmark test sets, whose data contain no exponentially growing outputs. Here REGSTABLE (ours) and GLOBSTABLE reach comparable accuracy. Since $\Theta_{\text{GLOBSTABLE}} \subseteq \Theta_{\text{REGSTABLE(ours)}}$, this indicates that REGSTABLE (ours) converges to a (nearly) globally stable model. Without diverging trajectories in the data, nothing in the loss counteracts the enlargement of the certificate (17), which drives the parameters towards the set certified by the standard sector conditions. The unconstrained models NOTSTABLE and LSTM are more accurate than both stable models. We attribute this gap to the conservatism of the stability constraints: conditions (12) are only sufficient, so the constrained parameter sets also exclude models that are stable on the data.

7 CONCLUSION

In this work, we proposed a novel definition of a regional stability certificate and derived parametric constraints for a class of recurrent neural networks. When the constraints are satisfied, the regional certificate is guaranteed to hold; this is enforced during training. We showed in two examples that our model learned the behavior of regionally stable systems, where globally stable models fail. On benchmarks without unbounded outputs, it matches the accuracy of globally stable models. Future work will address the conservatism of the conditions 12 and the two-stage training algorithm to keep the certificate close to the region covered by the data.

AI USE STATEMENT

In this work, we used generative AI tools for coding, running experiments on a slurm cluster and connecting the nonlinear benchmark datasets to our code. In addition we used it to polish our writing. We have not used generative AI tools for deriving the proofs of our theory implementing the parametric constraints in our models and literature review. We take responsibility for the final content of this work, including text, claims or artifacts produced with aid of generative AI.

REFERENCES

Alexandra Baier, Decky Aspandi, and Steffen Staab. ReLiNet: stable and explainable multistep prediction with recurrent linear parameter varying networks. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pp. 3461–3469, 2023.

Gerben I Beintema, Maarten Schoukens, and Roland Tóth. Continuous-time identification of dynamic state-space models by deep subspace encoding. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*, 2023a.

Gerben I Beintema, Maarten Schoukens, and Roland Tóth. Deep subspace encoders for nonlinear system identification. *Automatica*, 156:111210, 2023b.

Fabio Bonassi, Marcello Farina, Jing Xie, and Riccardo Scattolini. On recurrent neural networks for learning-based control: recent results and ideas for future developments. *Journal of Process Control*, 114:92–104, 2022.

Fabio Bonassi, Alessio La Bella, Marcello Farina, and Riccardo Scattolini. Nonlinear MPC design for incrementally ISS systems with application to GRU networks. *Automatica*, 159:111381, 2024.

Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

JM Gomes Da Silva and Sophie Tarbouriech. Antiwindup design with guaranteed regions of stability: an LMI-based approach. *IEEE Transactions on Automatic Control*, 50(1):106–111, 2005.

Mahyar Fazlyab, Alexander Robey, Hamed Hassani, Manfred Morari, and George Pappas. Efficient and accurate estimation of Lipschitz constants for deep neural networks. *Advances in Neural Information Processing Systems*, 32, 2019.

Daniel Frank, Tobias Holicki, Carsten W Scherer, and Steffen Staab. Regularised neural network-based nonlinear system identification with prior system knowledge. *International Journal of Control*, pp. 1–29, 2026.

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

Fangda Gu, He Yin, Laurent El Ghaoui, Murat Arcak, Peter Seiler, and Ming Jin. Recurrent neural network controllers synthesis with stability guarantees for partially observed systems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 5385–5394, 2022.

Zheyuan Hu, Nazanin Ahmadi Daryakenari, Qianli Shen, Kenji Kawaguchi, and George Em Karniadakis. State-space models are accurate and efficient neural operators for dynamical systems. *Neural Networks*, 197:108496, 2026. ISSN 0893-6080. doi: https://doi.org/10.1016/j.neunet.2025.108496.

Zhong-Ping Jiang and Yuan Wang. Input-to-state stability for discrete-time nonlinear systems. *Automatica*, 37(6):857–869, 2001.

Hassan K Khalil. *Nonlinear systems*. Prentice-Hall, Upper Saddle River, NJ, 3rd edition, 2002. ISBN 0130673897.

Alessio La Bella, Marcello Farina, William D'Amico, and Luca Zaccarian. Regional stability conditions for recurrent neural network-based control systems. *Automatica*, 174:112127, 2025.

L. Ljung. *System Identification: Theory for the User*. Pearson Education, 1998. ISBN 9780132440530.

Marco Massimetti, Luca Zaccarian, Tingshu Hu, and Andrew R. Teel. Linear discrete-time global and regional anti-windup: an LMI approach. *International Journal of control*, 82(12):2179–2192, 2009.

Nima Mohajerin and Steven L Waslander. Multistep prediction of dynamic systems with recurrent neural networks. *IEEE transactions on neural networks and learning systems*, 30(11):3370–3383, 2019.

Peter Overschee and Bart Moor. *Subspace identification for linear systems: Theory-Implementation-Applications*. Springer, 1996.

Patricia Pauli, Dennis Gramlich, Julian Berberich, and Frank Allgöwer. Linear systems with neural network nonlinearities: Improved stability analysis via acausal Zames-Falb multipliers. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 3611–3618. IEEE, 2021a.

Patricia Pauli, Anne Koch, Julian Berberich, Paul Kohler, and Frank Allgöwer. Training robust neural networks using Lipschitz bounds. *IEEE Control Systems Letters*, 6:121–126, 2021b.

Alexey Pavlov, Alexander Pogromsky, Nathan van de Wouw, and Henk Nijmeijer. Convergent dynamics, a tribute to Boris Pavlovich Demidovich. *Systems & Control Letters*, 52(3-4):257–261, 2004.

Gianluigi Pillonetto, Aleksandr Aravkin, Daniel Gedon, Lennart Ljung, Antonio H Ribeiro, and Thomas B Schön. Deep networks for system identification: a survey. *Automatica*, 171:111907, 2025.

Max Revay, Ruigang Wang, and Ian R Manchester. A convex parameterization of robust recurrent neural networks. *IEEE Control Systems Letters*, 5(4):1363–1368, 2020.

Max Revay, Ruigang Wang, and Ian R Manchester. Recurrent equilibrium networks: Unconstrained learning of stable and robust dynamical models. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 2282–2287. IEEE, 2021.

Shankar Sastry. *Nonlinear systems: analysis, stability, and control*, volume 10. Springer Science & Business Media, 2013.

Carsten Scherer and Siep Weiland. Linear matrix inequalities in control. *Lecture Notes, Dutch Institute for Systems and Control, Delft, The Netherlands*, 3(2), 2000.

Carsten Scherer, Pascal Gahinet, and Mahmoud Chilali. Multiobjective output-feedback control via LMI optimization. *IEEE Transactions on automatic control*, 42(7):896–911, 1997.

Carsten W Scherer. Robust exponential stability and invariance guarantees with general dynamic O'Shea-Zames-Falb multipliers. *arXiv preprint arXiv:2306.00571*, 2023.

Johan Schoukens and Lennart Ljung. Nonlinear system identification: A user-oriented road map. *IEEE Control Systems Magazine*, 39(6):28–99, 2019.

Maarten Schoukens. Improved initialization of state-space artificial neural networks. In *2021 European Control Conference (ECC)*, pp. 1913–1918. IEEE, 2021.

Maarten Schoukens, Max Champneys, Gerben Izaak Beintema, and Timothy James Rogers. Benchmarking for nonlinear system identification: Submission platform and baseline results. *Data-Centric Engineering*, 7:e40, 2026. doi: 10.1017/dce.2026.10067.

Mohammad Fahim Shakib, Alexander Yu Pogromsky, Alexey Pavlov, and Nathan van de Wouw. Computationally efficient identification of continuous-time Lur'e-type systems with stability guarantees. *Automatica*, 136:110012, 2022.

Mohammad Fahim Shakib, Roland Tóth, Alexander Yu Pogromsky, A Pavlov, and Nathan van de Wouw. Kernel-based learning of stable nonlinear state-space models. In *2023 62nd IEEE Conference on Decision and Control (CDC)*, pp. 2897–2902. IEEE, 2023.

Johan AK Suykens, Joos PL Vandewalle, and Bart L De Moor. *Artificial neural networks for modelling and control of non-linear systems*. Springer Science & Business Media, 1995.

Sophie Tarbouriech, Christophe Prieur, and João Manoel Gomes da Silva. Stability analysis and stabilization of systems presenting nested saturations. *IEEE Transactions on Automatic Control*, 51(8):1364–1371, 2006.

Sophie Tarbouriech, Germain Garcia, João Manoel Gomes da Silva Jr, and Isabelle Queinnec. *Stability and stabilization of linear systems with saturating actuators*. Springer Science & Business Media, 2011.

Wojciech Wawrzynski. Duffing-type oscillator under harmonic excitation with a variable value of excitation amplitude and time-dependent external disturbances. *Scientific reports*, 11(1):2889, 2021.

He Yin, Peter Seiler, and Murat Arcak. Stability analysis using quadratic constraints for systems with neural network controllers. *IEEE Transactions on Automatic Control*, 67(4):1980–1987, 2021.

Lotfi A Zadeh. From circuit theory to system theory. *Proceedings of the IRE*, 50(5):856–865, 1962.

A REGIONALLY STABLE SYSTEMS

This appendix illustrates the notions from Section 3 and the certificate of Section 4 on a one-dimensional model (Appendix A.1–A.3), shows the learned models on this system (Appendix A.4), and states the softening Duffing oscillator used in Section 6 (Appendix A.5).

A.1 A ONE-DIMENSIONAL REGIONALLY STABLE SYSTEM

Example 1 (One-dimensional system)**.** *Consider the one-dimensional system*

$$x_{k+1} = 2x_k - 1.1\,\mathrm{sat}(x_k) + u_k, \qquad \hat{y}_k = x_k, \tag{20}$$

where $\mathrm{sat}(\cdot)$ *is the saturation function. For* $u \equiv 0$ *the model has a stable equilibrium at* $x = 0$ *and two unstable equilibria at* $x = \pm 1.1$.

Analyzing (20) *by hand yields the regional certificate* $(\mathcal{X}, \mathcal{U}_{hand}, \mathcal{Y})$ *with*

$$\mathcal{X} = [-1, 1], \qquad \mathcal{U}_{hand}(x_k) = \{u_k \in \mathbb{R} \mid |u_k| \le 1 - 0.9|x_k|\}, \qquad \mathcal{Y} = [-1, 1],$$

which we call the hand-crafted *certificate. We simulate* (20) *in four cases, shown in Figure 2:*

Invariance and input-to-state stability *(Definitions 3 and 1). For* $x_0 = 0$, $u_0 = 0.4$ *and* $u_k = 0$ *for* $k > 0$, *the input is admissible, and the state stays in* $\mathcal{X}$ *and converges to zero.*

Sufficient, not necessary. *For* $x_0 = 0$, $u_0 = 1.09$ *and* $u_k = 0$ *for* $k > 0$, *the input violates* $\mathcal{U}_{hand}$, *yet the state stays bounded and converges to zero. The certificate is sufficient but not necessary for boundedness.*

Bounded input, unbounded output. *For* $x_0 = 0$, $u_0 = 1.2$ *and* $u_k = 0$ *for* $k > 0$, *the input violates* $\mathcal{U}_{hand}$ *and the state grows unbounded, although the input is bounded.*

Zero input, unbounded output. *For* $x_0 = 1.3 \notin \mathcal{X}$ *and* $u \equiv 0$, *the state grows unbounded without any input.*

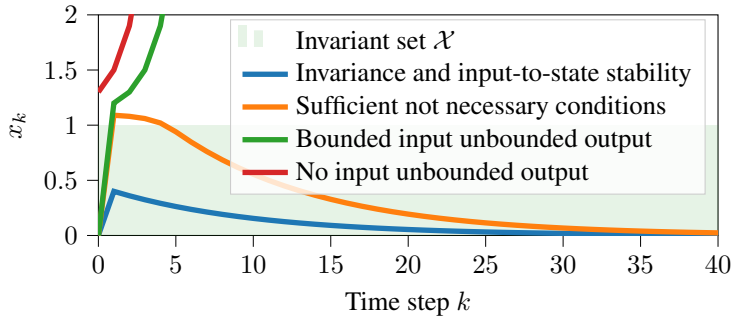


Figure 2: State trajectories of the one-dimensional model (20) for the four cases of Example 1. The shaded region is the invariant set $\mathcal{X} = [-1, 1]$ of the hand-crafted certificate.

Size of the hand-crafted input map. Problem 1 asks for model parameters θ and a regional certificate $(\mathcal{X}, \mathcal{U}, \mathcal{Y})$ with small prediction error and large admissible input map $\mathcal{U}$. The size functional (6) measures $\mathcal{U}$ by the largest input magnitude that is admissible for every $x_k \in \mathcal{X}$. For the hand-crafted certificate of Example 1,

$$\sigma_u(\mathcal{U}_{\text{hand}}) = \min_{x \in \mathcal{X}} (1 - 0.9|x|) = 0.1.$$

A.2 STANDARD AND GENERALIZED SECTOR CONDITIONS

Figure 3 shows the activation functions ψ discussed in Section 4.1.

The model of Example 1 shows why the standard sector conditions are too conservative for regionally stable models. The standard sector condition (8) admits the identity $\psi(v) = v$ as a worst case.

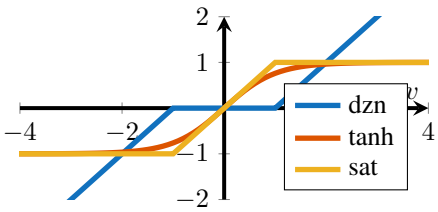


Figure 3: Activation functions $\tanh$, sat and dzn. The commonly used $\tanh$ does not tightly fill the standard sector (8). The deadzone dzn admits the generalized sector conditions (11).

With this nonlinearity the loop is unstable, so the model cannot be certified with standard sector conditions.

The situation changes with $H \neq 0$. On $\mathcal{L}(H) = [-1, 1]$ the argument of the deadzone satisfies $|v| \leq 1$, where $\mathrm{dzn}(v) = 0$. The worst case admitted by (11) is therefore an inactive nonlinearity, for which the loop reduces to $A = 0.9$ and is stable. The generalized sector condition thus certifies exactly the region on which the model is stable, whereas the standard sector condition must account for behavior that only occurs outside this region.

A.3 Regional certificate of the one-dimensional model

Example 2 (Continuation of Example 1)**.** *We now use the conditions of Theorem 1 to analyze model* (20)*. With* $\mathrm{sat}(v) = v - \mathrm{dzn}(v)$*, the model is of the form* (2) *with parameters*

$$\theta = \{A = 0.9,\ B = 1,\ B_2 = 1.1,\ C = 1,\ C_2 = 1,\ D_{21} = 0\}.$$

For fixed α*, solving the semidefinite program* (18) *yields* $s \approx 0.42$ *and* $P \approx 5.76$ *for* $\alpha = 0.99$*. The resulting certificate consists of*

- *the invariant state set* $\mathcal{X} = \mathcal{E}(P^{-1}/s^2) \approx [-0.99, 0.99]$*,*
- *the admissible input map* $\mathcal{U}_{SDP}(x_k) = \{u_k \in \mathbb{R} \mid u_k^2 \leq 0.42^2 - 0.99^2\, x_k^2/5.76\}$*, and*
- *the output set* $\mathcal{Y} = \mathcal{E}\big((s^2 C P C^{\mathrm{T}})^{-1}\big) \approx [-0.99, 0.99]$*, i.e.,* $\bar{y} \approx 0.99$*.*

By Lemma 3, $\sigma_u(\mathcal{U}_{SDP}) = s\sqrt{1-\alpha^2} \approx 0.059$*, compared with* $\sigma_u(\mathcal{U}_{hand}) = 0.1$*. The invariant sets of both certificates nearly coincide, but the input maps differ in shape. The quadratic Lyapunov function behind Theorem 1 yields an input bound that decreases quadratically in the state, whereas the hand-crafted bound decreases linearly.*

State-dependent versus static input constraint. The admissible input map $\mathcal{U}$ depends on the state. The largest state-independent input set it contains is the ball $\mathcal{U}_\delta := \mathcal{B}_\delta$ with $\delta = \sigma_u(\mathcal{U}_{\mathrm{SDP}})$, since $\mathcal{U}_\delta \subseteq \mathcal{U}_{\mathrm{SDP}}(x)$ for all $x \in \mathcal{X}$. To show that the state-dependent map is less conservative than this static constraint, we simulate model (20) with three input sequences: (i) $u \in \mathcal{U}_{\mathrm{SDP}}(x)$, (ii) $u \in \mathcal{U}_\delta$, and (iii) $u \notin \mathcal{U}_{\mathrm{SDP}}(x)$, which implies $u \notin \mathcal{U}_\delta$. The predictions are shown in Figure 4.

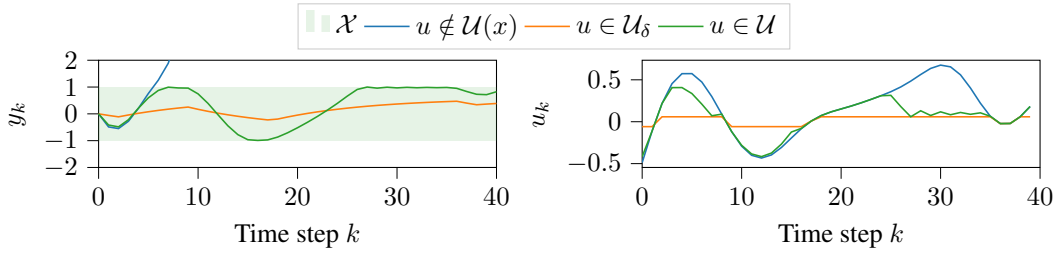
A.4 Learned models on the one-dimensional system

Figure 5 is the one-dimensional counterpart of Figure 1. It shows the regional stability tests described in Section 6.2 for the models trained on data from the one-dimensional system (20). Since $n = 1$, states are plotted over time instead of in the phase plane.

A.5 Softening Duffing oscillator

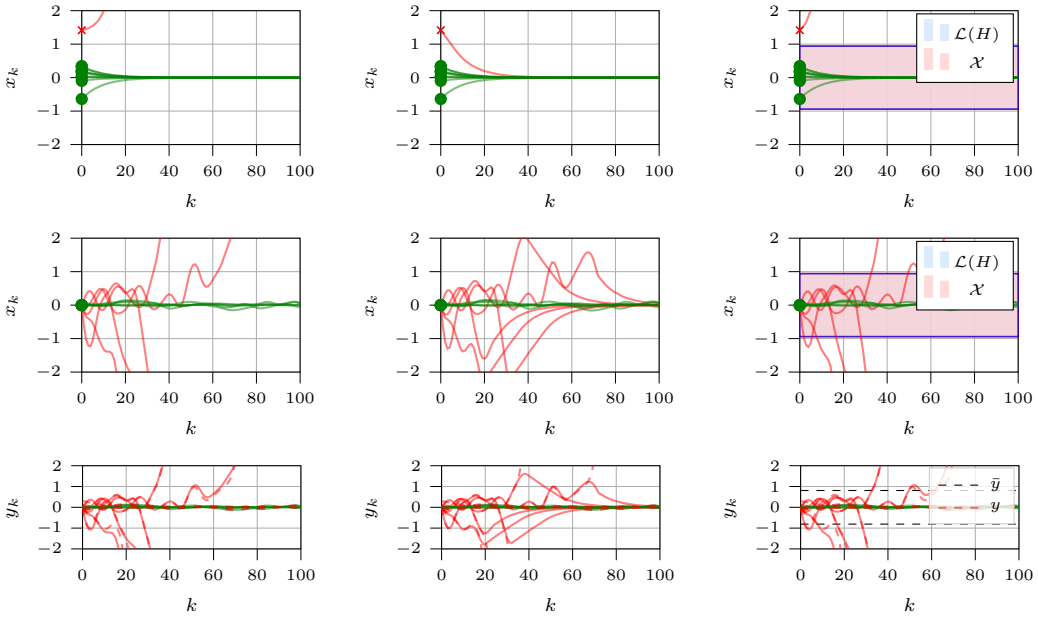
The softening Duffing oscillator is

$$\ddot{q} = -\delta_d\, \dot{q} - q + q^3 + u(t), \qquad \delta_d > 0, \tag{21}$$



(a) Output of model (20). (b) Inputs.

Figure 4: Certified and uncertified inputs for model (20). *Green:* an input from the admissible input map $\mathcal{U}$ drives the output across the entire output set $\mathcal{Y}$ without leaving it. *Orange:* an input from the static set $\mathcal{U}_\delta$ keeps the output away from the boundary of $\mathcal{Y}$, which shows that the static constraint is more conservative. *Blue:* an input outside $\mathcal{U}$ lets the output grow unbounded.

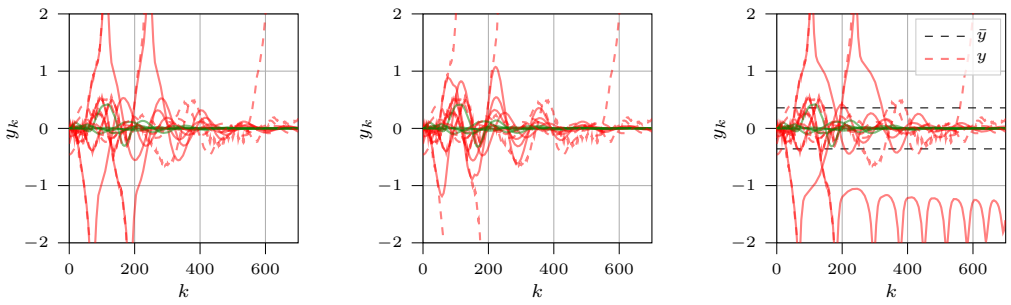


(a) NotStable (b) GlobStable (c) RegStable (ours)

Figure 5: Regional stability tests on the one-dimensional system. *Row 1* (initial state outside $\mathcal{X}$): state over time for $x_0 \neq 0$ and $u \equiv 0$. *Row 2* (input outside $\mathcal{U}$): state over time for ten rollouts from $x_0 = 0$, each driven by one test input. Red marks the five inputs for which the output of the data-generating system Σ diverges, green the five for which it converges. *Row 3*: output over time for the rollouts of row 2. Solid lines are model predictions, dashed red lines the output of Σ, truncated once it has diverged.

with damping $\delta_d = 0.3$. The cubic term has the opposite sign of the linear restoring term, so the oscillator is regionally but not globally stable. The model is discretized with step size 0.05 (Section 6).

Figure 6 shows the predicted outputs of the learned models over time, together with the output of (21). The phase-plane view of the same models is shown in Figure 1.



(a) NOTSTABLE (b) GLOBSTABLE (c) REGSTABLE (ours)

Figure 6: Predicted outputs of the learned models on the softening Duffing oscillator (21). Solid lines are model predictions: red for inputs under which the output of the oscillator grows exponentially, green for inputs under which it stays bounded. Dashed lines are the outputs of the oscillator.

B PROOFS AND REMARKS

B.1 PROOF OF LEMMA 1

Proof. Let $\mathcal{U}_1, \mathcal{U}_2$ be admissible input maps on $\mathcal{X}$ with $\mathcal{U}_1(x) \subseteq \mathcal{U}_2(x)$ for all $x \in \mathcal{X}$, and write $r_j(x) := \sup\{\delta \geq 0 : \mathcal{B}_\delta \subseteq \mathcal{U}_j(x)\}$, so that $\sigma_{\mathrm{u}}(\mathcal{U}_j) = \inf_{x\in\mathcal{X}} r_j(x)$. Fix $x \in \mathcal{X}$ and let $\delta \geq 0$ satisfy $\mathcal{B}_\delta \subseteq \mathcal{U}_1(x)$. By hypothesis $\mathcal{B}_\delta \subseteq \mathcal{U}_2(x)$, so every δ admissible in the supremum defining $r_1(x)$ is also admissible in the supremum defining $r_2(x)$; taking suprema gives $r_1(x) \leq r_2(x)$. Since $x \in \mathcal{X}$ was arbitrary, $\inf_{x\in\mathcal{X}} r_1(x) \leq \inf_{x\in\mathcal{X}} r_2(x)$, that is $\sigma_{\mathrm{u}}(\mathcal{U}_1) \leq \sigma_{\mathrm{u}}(\mathcal{U}_2)$. □

B.2 PROOF OF LEMMA 2

Proof. Recall that we have $\psi = \mathrm{dzn}$, since Δ is decentralized, it suffices to show for all $i \in \{1, \ldots, q\}$ and $\Lambda = \mathrm{diag}(\lambda^1, \ldots, \lambda^q)$ that the condition

$$\begin{pmatrix} \psi(v^i) \\ v^i + h^i x \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} -2\lambda^i & \lambda^i \\ \lambda^i & 0 \end{pmatrix} \begin{pmatrix} \psi(v^i) \\ v^i + h^i x \end{pmatrix} \geq 0 \tag{22}$$

is satisfied for any $v^i \in \mathbb{R}$ and any $x \in \mathbb{R}^n$ such that $|h^i x| \leq 1$. When expanding (22) we get the scalar inequality

$$2\lambda^i \left(-\psi(v^i) + v^i + x^{\mathrm{T}}(h^i)^{\mathrm{T}}\right) \psi(v^i) \geq 0. \tag{23}$$

We check all cases of the nonlinear function ψ:

- Assume $|v^i| \leq 1$, then $\psi(v^i) = 0$, hence (23) is satisfied.
- Assume $v^i > 1$, then (23) reads as follows:

  $$2\lambda^i \left(1 + x^{\mathrm{T}}(h^i)^{\mathrm{T}}\right) (v^i - 1) \geq 0. \tag{24}$$

  Since $x \in \mathcal{L}(H)$, we have $x^{\mathrm{T}}(h^i)^{\mathrm{T}} \in [-1, 1]$ and, hence, $(1 + x^{\mathrm{T}}(h^i)^{\mathrm{T}}) \geq 0$. Using the fact that $\lambda^i > 0$, we conclude that (24) (and hence (23) is satisfied).
- Assume $v^i < -1$, then (23) reads as follows:

  $$2\lambda^i \left(-1 + x^{\mathrm{T}}(h^i)^{\mathrm{T}}\right) (v^i + 1) \geq 0. \tag{25}$$

  Using the same arguments as above, we see that $(-1 + x^{\mathrm{T}}(h^i)^{\mathrm{T}}) \leq 0$ and $(v^i + 1) < 0$. Then, since $\lambda^i > 0$, (25) and (23) are satisfied.

Since i was arbitrary, we conclude that $\Gamma(v, x, \Lambda, H) \geq 0$ for all $v \in \mathbb{R}^q$ and all $x \in \mathcal{L}(H)$, which completes the proof. □

B.3 PROOF OF THEOREM 1

Proof. We start by performing a congruence transformation on the matrices in conditions (12), i.e., $\mathrm{diag}(X, I, \Lambda, X)$ on (12a) and $\mathrm{diag}(I, X)$ on (12b) leads to

$$\left(\begin{array}{ccc|c} -\alpha^2 XPX & 0 & XPC_2^{\mathrm{T}}\Lambda + XL^{\mathrm{T}}\Lambda & XPA^{\mathrm{T}}X \\ \star & -I & D_{21}^{\mathrm{T}}\Lambda & B^{\mathrm{T}}X \\ \star & \star & -2\Lambda M\Lambda & \Lambda M B_2^{\mathrm{T}}X \\ \hline \star & \star & \star & -XPX \end{array}\right) \preceq 0$$

and

$$\begin{pmatrix} 1/s^2 & l^i X \\ X(l^i)^{\mathrm{T}} & XPX \end{pmatrix} \succeq 0 \qquad \text{for all} \qquad i = 1, \ldots, q.$$

Now, we substitute $P = X^{-1}$, $M = \Lambda^{-1}$, and $L = HP$ (note that $l^i X = h^i$) to get

$$\left(\begin{array}{ccc|c} -\alpha^2 X & 0 & (C_2^{\mathrm{T}} + H^{\mathrm{T}})\Lambda & A^{\mathrm{T}}X \\ \star & -I & D_{21}^{\mathrm{T}}\Lambda & B^{\mathrm{T}}X \\ \star & \star & -2\Lambda & B_2^{\mathrm{T}}X \\ \hline \star & \star & \star & -X \end{array}\right) \preceq 0 \tag{27a}$$

and

$$\begin{pmatrix} 1/s^2 & h^i \\ (h^i)^{\mathrm{T}} & X \end{pmatrix} \succeq 0 \qquad \text{for all} \qquad i = 1, \dots, q. \tag{27b}$$

By taking the *Schur*-complement of (27a) with respect to the last $-X$ block we get

$$-\begin{pmatrix} A^{\mathrm{T}}X \\ B^{\mathrm{T}}X \\ B_2^{\mathrm{T}}X \end{pmatrix} (-X)^{-1} \begin{pmatrix} XA & XB & XB_2 \end{pmatrix} + \begin{pmatrix} -\alpha^2 X & 0 & (C_2^{\mathrm{T}} + H^{\mathrm{T}})\Lambda \\ \star & -I & D_{21}^{\mathrm{T}}\Lambda \\ \star & \star & -2\Lambda \end{pmatrix} \preceq 0, \tag{28}$$

which can be rewritten as

$$\begin{aligned}(\star)^{\mathrm{T}} \begin{pmatrix} -\alpha^2 X & 0 \\ 0 & X \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ A & B & B_2 \end{pmatrix} - (\star)^{\mathrm{T}} \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & I & 0 \\ C & D & D_{12} \end{pmatrix} \\ + (\star)^{\mathrm{T}} \begin{pmatrix} -2\Lambda & \Lambda \\ \Lambda & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & I \\ C_2 + H & D_{21} & 0 \end{pmatrix} \preceq 0. \end{aligned} \tag{29}$$

This is a condition that appears in robust control literature, see for example Scherer & Weiland (2000).

When we multiply from left with $\begin{pmatrix} x_k^{\mathrm{T}} & u_k^{\mathrm{T}} & w_k^{\mathrm{T}} \end{pmatrix}$ and from right $\begin{pmatrix} x_k^{\mathrm{T}} & u_k^{\mathrm{T}} & w_k^{\mathrm{T}} \end{pmatrix}^{\mathrm{T}}$ and use $v_k = C_2 x_k + D_{21} u_k$ from (2) we get

$$\begin{aligned}\begin{pmatrix} x_k \\ x_{k+1} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} -\alpha^2 X & 0 \\ 0 & X \end{pmatrix} \begin{pmatrix} x_k \\ x_{k+1} \end{pmatrix} - \begin{pmatrix} u_k \\ y_k \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_k \\ y_k \end{pmatrix} \\ + \begin{pmatrix} w_k \\ v_k + Hx_k \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} -2\Lambda & \Lambda \\ \Lambda & 0 \end{pmatrix} \begin{pmatrix} w_k \\ v_k + Hx_k \end{pmatrix} \leq 0, \end{aligned} \tag{30}$$

for all $\begin{pmatrix} x_k^{\mathrm{T}} & u_k^{\mathrm{T}} & w_k^{\mathrm{T}} \end{pmatrix}^{\mathrm{T}} \neq 0$. Introducing a quadratic *Lyapunov* function $V(x) = x^T X x$, we rewrite (30) into

$$V(x_{k+1}) - \alpha^2 V(x_k) - (u_k)^{\mathrm{T}} u_k + \begin{pmatrix} w_k \\ v_k + Hx_k \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} -2\Lambda & \Lambda \\ \Lambda & 0 \end{pmatrix} \begin{pmatrix} w_k \\ v_k + Hx_k \end{pmatrix} \leq 0. \tag{31}$$

Let us now assume that $x_k \in \mathcal{L}(H)$ (which will be proved in the next step), which, by Lemma 2, implies that $\Gamma(v, x, \Lambda, H)$ is non-negative (for $w_k = \Delta(v_k)$) and thus we can drop the last term of inequality (31) which yields:

$$V(x_{k+1}) - \alpha^2 V(x_k) - (u_k)^{\mathrm{T}} u_k \leq 0. \tag{32}$$

Now we show that the ellipsoid $\mathcal{E}(X/s^2)$ is forward invariant and $\mathcal{E}(X/s^2) \subset \mathcal{L}(H)$. Assume that for some k, $V(x^k) \leq s^2$. Rewriting (32) and using the condition on the input size $(u_k)^{\mathrm{T}} u_k \leq s^2 - \alpha^2 V(x_k)$, we get

$$V(x_{k+1}) \leq \alpha^2 V(x_k) + \underbrace{(u_k)^{\mathrm{T}} u_k}_{\leq s^2 - \alpha^2 V(x_k)} \leq \alpha^2 V(x_k) + s^2 - \alpha^2 V(x_k) = s^2. \tag{33}$$

Since the initial state is assumed to satisfy $V(x_0) < s^2$, we conclude that for every $k \geq 0$ it holds that $\frac{1}{s^2}(x_k)^{\mathrm{T}} X x_k \leq 1$ and thus $x_k \in \mathcal{E}(X/s^2)$, i.e. $\mathcal{E}(X/s^2)$ is forward invariant under input from $\mathcal{U}$.

To show that $\mathcal{E}(X/s^2) \subset \mathcal{L}(H)$, we pick an $x_\star \in \mathcal{E}(X/s^2)$, i.e., $\frac{1}{s^2}(x_\star)^{\mathrm{T}} X x_\star \leq 1$. Applying the *Schur*-complement to (27b) gives $s^2 (h^i)^{\mathrm{T}} h^i \preceq X$, which implies

$$(x_\star)^{\mathrm{T}} (h^i)^{\mathrm{T}} h^i x_\star \leq \frac{1}{s^2} (x_\star)^{\top} X x_\star \leq 1, \tag{34}$$

for all $i = 1, \dots, q$. To be in the set $\mathcal{L}(H)$, $x_\star$ needs to satisfy $\|Hx_\star\|_\infty \leq 1$. Note that

$$\|Hx_\star\|_\infty = \sup_i |h^i x_\star| = \sup_i \sqrt{(x_\star)^{\mathrm{T}} (h^i)^{\mathrm{T}} h^i x_\star}. \tag{35}$$

Since (34) holds for all i, we therefore conclude that $\|Hx_\star\|_\infty \leq 1$ and hence, $x_\star \in \mathcal{L}(H)$. Furthermore, since $x_\star$ was arbitrary, we conclude that all points in $\mathcal{E}(X/s^2)$ are also included in $\mathcal{L}(H)$, hence $\mathcal{E}(X/s^2) \subset \mathcal{L}(H)$.

Now we have proven that the set $\mathcal{X}$ is forward invariant under the dynamics of model (2) with inputs satisfying (13) and $x_0 \in \mathcal{E}(X/s^2)$.

To finally show the bound (14), we unroll (32) to get

$$\begin{aligned} V(x_k) &\leq \alpha^{2k} V(x_0) + \sum_{\tau=0}^{k-1} \alpha^{2(k-1-\tau)} \underbrace{(u_\tau)^{\mathrm{T}} u_\tau}_{\leq s^2(1-\alpha^2)} \\ &\leq \alpha^{2k} V(x_0) + \frac{1}{1-\alpha^2} \sup_{0\leq\tau<k} \|u_\tau\|^2. \end{aligned} \tag{36}$$

by exploiting the finite geometric sum $\sum_{\tau=0}^{k-1} \alpha^\tau \leq \frac{1}{1-\alpha^2}$. Using the bounds $\lambda_{\min}|x|^2 \leq V(x) \leq \lambda_{\max}|x|^2$ and taking the square root of 36, we get

$$|x_k| \leq \min\Big\{ \sqrt{\tfrac{\lambda_{\max}}{\lambda_{\min}}}\, \alpha^k |x_0| + \frac{\sup_{0\leq j<k} |u_j|}{\sqrt{\lambda_{\min}(1-\alpha^2)}}, \frac{s}{\sqrt{\lambda_{\min}}} \Big\}. \tag{37}$$

here we exploit $\sqrt{(a+b)} \leq \sqrt{a} + \sqrt{b}$ for $a, b \geq 0$ to split the square root of the sum into a sum of square roots, which is a valid upper bound. This leads to the bound (14).

To show the output bound we need to ensure that whenever $x_k \in \mathcal{X}$ it holds that $\hat{y}_k \in \mathcal{Y}$. Let us start with the claim $\hat{y}_k \in \mathcal{Y}$ and show that it implies $x_k \in \mathcal{X}$ which means

$$\hat{y}_k^{\mathrm{T}} (s^2 CPC^{\mathrm{T}})^{-1} \hat{y}_k \leq 1$$

substituting $\hat{y}_k = Cx_k$ gives

$$x_k^{\mathrm{T}} C^{\mathrm{T}} (CPC^{\mathrm{T}})^{-1} Cx_k \leq s^2.$$

Since C is not required to be square and invertible, it does not hold in general that $C^{\mathrm{T}}(CPC^{\mathrm{T}})^{-1}C = P^{-1}$. Let us define

$$a_k := x_k^{\mathrm{T}} C^{\mathrm{T}} (CPC^{\mathrm{T}})^{-1} Cx_k, \qquad w_k := (CPC^{\mathrm{T}})^{-1} Cx_k$$

where CPC^{T} is invertible, since $P \succ 0$ and C has full row rank. Then $a_k = x_k^{\mathrm{T}} C^{\mathrm{T}} w_k$ and $w_k^{\mathrm{T}} CPC^{\mathrm{T}} w_k = a_k$. Let $P^{1/2}$ denote the symmetric square root of P. Inserting $I = P^{-1/2}P^{1/2}$ and applying Cauchy-Schwarz inequality yields

$$a_k = (P^{-1/2}x_k)^{\mathrm{T}} (P^{1/2} C^{\mathrm{T}} w_k) \leq |P^{-1/2}x_k|\, |P^{1/2} C^{\mathrm{T}} w_k| = \sqrt{V(x_k)}\, \sqrt{a_k}.$$

If $a_k = 0$, the claim holds trivially. Otherwise, dividing by $\sqrt{a_k}$ and squaring gives $a_k \leq V(x_k) \leq s^2$, which is the required inequality. Hence $\hat{y}_k \in \mathcal{Y}$ whenever $x_k \in \mathcal{X}$, and since $\mathcal{X}$ is forward invariant, $\hat{y}_k \in \mathcal{Y}$ for all $k \geq 0$, which finishes the proof. □

B.4 Proof of Lemma 3

Proof. For each $x \in \mathcal{X}$ the set $\mathcal{U}(x)$ is the ball of radius $\rho(x) := \sqrt{s^2 - \alpha^2 V(x)}$ centred at the origin, hence $\mathcal{B}_\delta \subseteq \mathcal{U}(x)$ if and only if $\delta \leq \rho(x)$, and the inner supremum in (6) equals $\rho(x)$. Since ρ is strictly decreasing in $V(x)$ and $0 \leq V(x) \leq s^2$ on $\mathcal{X}$, the infimum is attained on the boundary $V(x) = s^2$, giving $\sigma_{\mathrm{u}}(\mathcal{U}) = \sqrt{s^2 - \alpha^2 s^2} = s\sqrt{1-\alpha^2}$. In addition, this proves that the admissible input map (13) is uniformly bounded with $\bar{u} = s$. □

B.5 Remark on bounded inputs for RNN

Remark 3. *Let $\mathcal{X}$ be forward invariant for $\mathcal{S}(\theta)$ with respect to $\mathcal{U}$ and let $x_0 \in \mathcal{X}$. By induction, every input sequence $u_k \in \mathcal{U}(x_k)$ generates a state $x_k \in \mathcal{X}$ for all $k \geq 0$. Uniform boundedness of $\mathcal{U}$ then give $|u_k| \leq \bar{u}$ for all $k \geq 0$, i.e., $\|u\|_\infty \leq \bar{u}$ and $u \in \ell_\infty^r$. Hence, the ISS bound* (4) *is well defined for all admissible inputs.*

C TRAINING AND ALGORITHM

This appendix details the training procedure of Section 5: how trajectories are initialized, how feasibility of (12) is obtained and maintained, when the admissible input map is enlarged, and how the learning rate and barrier weight are scheduled. Algorithm 1 summarizes the procedure.

C.1 TRAINING DETAILS

Washout initialization. The dataset (1) contains inputs and outputs but no state measurements, so the initial state $x_0^{(i)}$ required to simulate $\mathcal{S}(\theta)$ on trajectory i is unknown. We set $x_0^{(i)} = 0$ and discard the first w predictions of each trajectory from the loss, which lets the effect of the unknown initial state decay before the residuals are evaluated (Beintema et al., 2023b; Mohajerin & Waslander, 2019; Baier et al., 2023; Schoukens, 2021). The simulation error in (17) is therefore evaluated over $k = w, \dots, N_i - 1$.

Initial feasibility. The barrier terms in (17) are only defined when the conditions (12) are satisfied. Training therefore has to start from a feasible point. If the initial parameters are infeasible, we solve a feasibility problem before the first gradient step (Algorithm 1, lines 4–6).

Restoring feasibility. The barrier enforces the constraints only softly, so a gradient step can cross the boundary of the feasible set, where the loss is undefined. After each update we therefore test whether (θ, κ) still satisfies (12). If it does not, we freeze θ and α and solve a feasibility semidefinite program for the certificate variables alone, first at the current s and, if that fails, with s as an additional optimization variable. Both programs are convex. If no feasible κ exists for the updated θ, the step is discarded and the previous parameters are restored, which amounts to skipping that training step.

Trigger for enlarging the admissible input map. Solving (18) after every gradient step is computationally expensive. We therefore solve it only at the end of an epoch, and only when the certificate has become too small to admit the training inputs, that is, when

$$c_{\max} := \max_{i,k} \left(|u_k^{(i)}|^2 - s^2 + \alpha^2 V(x_k^{(i)}) \right) > 0. \tag{38}$$

The test (38) is a computational trigger, not a constraint of the learning problem. The objective remains $\bar{\sigma}$, and admissibility of the training inputs only serves as an inexpensive indicator that s has drifted away from its optimum. The certificate we report for a trained model is obtained by solving (18) once for the epoch with the lowest validation error.

Scheduling. When the validation loss has not improved for a fixed number of epochs, the learning rate η and the barrier weight ν are both decreased by a constant factor $\gamma \in (0, 1)$. Decreasing ν progressively reduces the influence of the barrier and allows the parameters to approach the boundary of the feasible set, which mirrors the continuation used in interior-point methods (Boyd et al., 2004).

C.2 ALGORITHM

Algorithm 1 gives the pseudocode of our approach to Problem 1.

Algorithm 1 Certified training of the Lur'e model (2). The dynamics θ and the certificate $\kappa = (P, M, L, \hat{s}, \tau)$ are learned jointly by gradient descent; two convex programs restore (12) whenever a step leaves the feasible set.

Require: dataset $\mathcal{D}$, barrier weight ν, step size η, decay $\gamma \in (0,1)$
$\Theta_{\mathrm{f}} := \{(\theta, \kappa, \alpha) \mid \mathbf{F} \prec 0,\ \mathbf{G}_i \succ 0,\ i = 1, \dots, q\}$, the strictly feasible set of (12)
$c_k := (u_k)^{\mathrm{T}} u_k - s^2 + \alpha^2 V(x_k)$, so (13) holds iff $c_k \le 0$

Initialization
$A \leftarrow I + \delta t A_c$, A_c Hurwitz companion with random last row; B, C from the data scaling
$B_2, C_2, D_{21} \sim \mathcal{N}(0, \sigma^2)$ ▷ Δ is not identifiable a priori
$(P, M, L, s) \leftarrow (I, I, 0, 1), \quad \alpha \leftarrow 1/(1 + e^{-\tau_0})$
if $(\theta, \kappa, \alpha) \notin \Theta_{\mathrm{f}}$ **then** ▷ a blind D_{21} is generically infeasible
    $(\kappa, D_{21}) \leftarrow \textsc{MaxS}(\theta, \alpha; \{D_{21}\})$, **abort** if infeasible
end if

Training
for epoch $= 1, \dots, N_{\mathrm{ep}}$ **do**
    for minibatch $\mathcal{B} \subset \mathcal{D}$ **do**
        $(\theta^-, \kappa^-, \tau^-) \leftarrow (\theta, \kappa, \tau)$ ▷ the step is provisional
        $(\theta, \kappa, \tau) \leftarrow (\theta, \kappa, \tau) - \eta \nabla \mathbf{L}(\mathcal{B})$ ▷ $\mathbf{L}$ as in (17)
        if $(\theta, \kappa, \alpha) \notin \Theta_{\mathrm{f}}$ **then**
            $\kappa \leftarrow \textsc{Feas}(\theta, \alpha, s)$ ▷ repair at the learned scale
            if infeasible **then** $\kappa \leftarrow \textsc{Feas}(\theta, \alpha)$ ▷ repair with s released
            if infeasible **then** $(\theta, \kappa, \tau) \leftarrow (\theta^-, \kappa^-, \tau^-)$ ▷ roll the step back
        end if
    end for
    roll out $\mathcal{D}$ and evaluate $c_{\max} = \max_{i,k} c_k^{(i)}$
    if $c_{\max} > 0$ **then** ▷ $\mathcal{X}$ no longer admits the training inputs
        $\kappa \leftarrow \textsc{MaxS}(\theta, \alpha; \emptyset)$ ▷ the *only* step that enlarges s
    end if
    if plateau **then** $\eta \leftarrow \gamma\eta$, $\nu \leftarrow \gamma\nu$ ▷ barrier continuation
end for

θ and α enter both programs as constants, so both are SDPs; $\hat{s} := 1/s^2$ makes (12b) linear.
function $\textsc{MaxS}(\theta, \alpha; \mathcal{V})$
    return $\arg\min \hat{s}$ over $(P, M, L, \hat{s}) \cup \mathcal{V}$ subject to (12)
end function
function $\textsc{Feas}(\theta, \alpha[, s])$
    return any (P, M, L) satisfying (12), at the given s or with $\hat{s}$ free
end function

D Experiments

D.1 Experimental setup

All models are trained on a GPU cluster (NVIDIA A100). For each model and dataset, we perform three runs with identical hyperparameters and different random seeds, and report the lowest NRMSE on the held-out test set. The semidefinite programs (18) and the feasibility problems of Appendix C are solved with MOSEK.

D.2 Output certificate

Table 1 reports the certified output bound $\bar{y}$ from Section 4.2 together with $y_{\max}$, the largest absolute output among the bounded trajectories in the training data.

For the one-dimensional system, $y_{\max} \le \bar{y}$, so the certified output set covers all bounded training data.

For the softening Duffing oscillator, the certified output set does not cover all bounded training data. This is expected, since the conditions (12) are only sufficient: an input that is not admissible does not necessarily lead to an unbounded output.

D.3 HYPERPARAMETERS

Table 3 lists the hyperparameters of the models reported in Tables 1 and 2. The columns give the number of learnable parameters n_θ, the training time, the washout length w (Warmup), the batch size, the number of nonlinearities n_w, and the barrier weight ν in (17). The barrier weight does not apply to the unconstrained models NOTSTABLE and LSTM.

Table 3: Hyperparameters of the learned models.

Dataset	Model	n_θ	Time	Warmup	Batch	n_w	ν
OneDExample	NOTSTABLE	32	58.8 min	50	32	4	–
	GLOBSTABLE	32	1.5 min	50	32	4	10^{-3}
	LSTM	12961	15.1 min	50	32	32	–
	REGSTABLE (ours)	37	2.0 h	50	32	4	10^{-3}
Soft. Duffing osc.	NOTSTABLE	782	5.9 h	500	32	24	–
	GLOBSTABLE	732	1.9 h	500	32	24	10^{-3}
	LSTM	12961	23.8 min	500	32	32	–
	REGSTABLE (ours)	782	6.4 h	500	32	24	10^{-3}
Silverbox	NOTSTABLE	399	3.4 h	50	8	16	–
	GLOBSTABLE	365	2.7 h	50	8	16	10^{-5}
	LSTM	331393	2.4 h	50	8	128	–
	REGSTABLE (ours)	399	3.4 h	50	8	16	10^{-5}
EMPS	NOTSTABLE	399	6.2 h	1000	8	16	–
	GLOBSTABLE	365	5.3 h	1000	8	16	10^{-4}
	LSTM	199297	4.9 h	1000	8	128	–
	REGSTABLE (ours)	399	6.1 h	1000	8	16	10^{-4}
CED	NOTSTABLE	303	2.5 h	50	4	12	–
	GLOBSTABLE	265	2.2 h	50	4	12	10^{-4}
	LSTM	331393	33 s	50	8	128	–
	REGSTABLE (ours)	303	2.6 h	50	4	12	10^{-4}
F16	NOTSTABLE	677	6.0 h	500	8	16	–
	GLOBSTABLE	579	5.5 h	500	8	16	10^{-3}
	LSTM	331651	2.9 h	50	8	128	–
	REGSTABLE (ours)	677	6.3 h	500	8	16	10^{-3}
WH System	NOTSTABLE	887	2.4 h	50	8	20	–
	GLOBSTABLE	765	2.0 h	50	8	20	10^{-4}
	LSTM	331393	3.8 h	50	8	128	–
	REGSTABLE (ours)	887	2.7 h	50	8	20	10^{-4}