Model predictive control design for dynamical systems learned by Long Short-Term Memory Networks

Enrico Terzi¹, Marcello Farina¹, and Riccardo Scattolini¹

Abstract—In this paper the stability-related properties of Long Short-Term Memory (LSTM) networks are analyzed, and their use as the model of the plant in the design of Model Predictive Controllers (MPC) is investigated. First, sufficient conditions guaranteeing Incremental Input-to-State stability (δISS) of LSTM are derived, and it is shown that this property can be enforced during the training of the network. Then, the design of an observer with guaranteed convergence of the state estimate to the true one is addressed, the observer is then embedded in a MPC scheme designed for the solution of the tracking problem. The resulting closed-loop scheme is proved to be asymptotically stable. The training algorithm and control scheme are tested numerically on the simulator of a pH reactor, the reported results confirm the effectiveness of the proposed approach.

I. INTRODUCTION

The availability of large and informative dataset, collected on plants during long periods of time and spanning many different working conditions, is nowadays a typical starting point in control-related projects [24], [19]. Also thanks to the recent introduction and popularity of novel tools and algorithms for extracting information from data [41], engineers and scientists are increasingly focusing on data-based identification and control techniques [5], [3]. Several approaches are aimed at the direct learning of the controller from data [39], these algorithms can be either based on a possibly reference - model, like in the Virtual Reference Feedback Tuning approach [6] and in Iterative Learning [5], or exploit model-free techniques, like Reinforcement Learning [30]. On the other hand, indirect approaches are aimed at first finding a model of the plant, based on which the controller is designed. In the latter category, a quite recent model class that has gained extraordinary attention and popularity is represented by Neural Networks (NN) [20], which have proven to be effective in a large variety of contexts and tasks, like image [28], speech [15] and handwriting recognition [16], prediction [43], and forecasting [44], [21].

In the control context, and in order to account for the dynamic nature of the systems to be controlled, recurrent Neural Networks (rNN) have already been studied [33], [10] and used in a number of applications [40], [31], [26], [29]. In rNN, the output of the network is fed-back as input, so constituting a loop which allows to properly describe the dynamics of the system. However, the tuning of rNN calls for a complex training algorithm, that is affected by the so-called “vanishing (or exploding) gradient” problem [22]. Essentially, this prevents a proper training given the recursive equations featuring the network, that cause a vanishing (or explosion) of information and gradient over the iterations. Up to date, only a couple of architectures proved to be able to practically overcome this issue, namely Echo State (ESN) [25] and Long Short Term Memory (LSTM) networks [17].

Although ESN have proven to be effective and characterized by a simple training procedure, LSTM [13] are gaining a wider popularity [13]. First introduced in 1997 [23], LSTM are nowadays widely used for several tasks [42],[38] and in everyday’s devices, such as mobile phones and GPS navigators for speech recognition. This diffusion is due to their flexibility and ability to recover long-term dependencies across the data thanks to their internal states. In the context of dynamical systems and control, some very recent stability results about their autonomous, i.e. non forced, version have been described in [9], [8], and [1], where also an analysis of their equilibria has been reported.

In this paper, we investigate the use of LSTM in the context of Model Predictive Control, on the same line of research as [4],[35]. First, we establish theoretically-sound conditions on their parameters (internal weights) guaranteeing the Incremental Input-to-State Stability property (δISS [11]); notably these conditions explicitly depend on the model parameters and can be forced in the training phase of the network. Then, assuming that the trained net exactly represents the model of the system and relying on δISS, we design an observer guaranteeing that the estimated state asymptotically converges to the true value. Finally, based on the LSTM model and on the state observer, we design a MPC control algorithm solving the tracking problem for constant state and input reference values and in presence of input constraints. The stability of the equilibrium point is obtained with a suitable tuning of the MPC design parameters, i.e. the state and control weighting functions and the terminal cost.
Notably, no terminal constraints are required, which makes easier the tuning procedure.

The performances of the overall control system are tested numerically on the simulator of a pH neutralization process [18], that represents a well-known benchmark for nonlinear SISO systems. The modeling capability of the trained LSTM is first quantitatively evaluated on a validation dataset, then closed-loop experiments are reported, witnessing the potentialities of the proposed approach.

The paper is organized as follows: in Section II the dynamic model of LSTM is analysed and its δISS property is established. In Section III the design of the stabilizing observer and of the MPC algorithm is discussed, while in Section IV the numerical example is described. Finally, conclusions and hints for future work are included in Section V. An Appendix reports the proofs of the theoretical results.

Notation and basic definitions.

We denote $v(j)$ the $j$-th entry of vector $v$. $0_{a,b}$ is the null matrix of dimension $a \times b$. $I_n$ is the identity matrix of order $n$. Moreover, given a vector $v$, we denote $\|v\|$ as the 2-norm of $v$, $\|v\|_2 = v^T A v$ its squared norm weighted by matrix $A$, and with $\|v\|_\infty$, its infinity norm, i.e., $\|v\|_\infty = \max_{j=1,\ldots,n} |v(j)|$, being $n$ the number of entries of $v$. $v^T$ denotes vector transpose and $\text{diag}(v)$ the diagonal matrix with $v$ on the diagonal. We denote with $|A|$ and with $\|A\|_m$ the induced 2-norm and $\|A\|_\infty$-norm, respectively, of $A$, while $\rho(M)$ is the spectral radius of the square matrix $M$ (i.e. maximum absolute value of its eigenvalues). Given an interval $[a,b] \subset \mathbb{R}$ and a positive integer $n$ we denote $[a,b]^n = \{x \in \mathbb{R}^n : x(j) \in [a,b], \forall j = 1, \ldots, n\}$. The same notation is applied for open intervals.

Consider the discrete-time system

$$\chi^+ = \varphi(\chi, u)$$

where $\chi$ is the state vector, $u$ is the input vector, and $\varphi(\cdot)$ is a nonlinear function of the input and the state. $\chi^+$ indicates the value of $\chi$ at the next step time. We indicate with $\chi_0(k)$ the solution to system (1) at time step $k$ starting from the initial state $\chi_0$ with input sequence $u_0(0), \ldots, u_0(k-1)$. Time index $k$ will be omitted where possible and clear from the context for better readability.

Now, we recall some notions, see also [11], useful for the following developments.

**Definition 1 ($\mathcal{K}$-Function):** A continuous function $\alpha : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is a class $\mathcal{K}$ function if $\alpha(s) > 0$ for all $s > 0$, it is strictly increasing, and $\alpha(0) = 0$.

**Definition 2 ($\mathcal{K}_*\alpha$-Function):** A continuous function $\alpha : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is a class $\mathcal{K}_*\alpha$ function if it is a class $\mathcal{K}$ function and $\alpha(s) \to \infty$ for $s \to \infty$.

**Definition 3 ($\mathcal{KL}^\mathcal{K}$-Function):** A continuous function $\beta : \mathbb{R}_{>0} \times \mathbb{Z}_{>0} \to \mathbb{R}_{>0}$ is a class $\mathcal{KL}^\mathcal{K}$ function if $\beta(s,k)$ is a class $\mathcal{K}$ function with respect to $s$ for all $k$, it is strictly decreasing in $k$ for all $s > 0$, and $\beta(s,k) \to 0$ as $k \to \infty$ for all $s > 0$.

**Definition 4 ($\delta$ISS[11]):** System (1) is called incrementally input-to-state stable in $\mathcal{X}$ with respect to $\mathcal{Y}$, if there exists a function $\beta \in \mathcal{KL}^\mathcal{K}$ and a function $\gamma \in \mathcal{K}_*\alpha$ such that for any $k \in \mathbb{Z}_{>0}$, any initial states $\chi_0, \chi_2 \in \mathcal{X}$, and any pair of disturbance sequences $u_1(0), u_1(1), \ldots, u_2(0), u_2(1), \ldots$, it holds that:

$$\|\chi_1(k) - \chi_2(k)\| \leq \beta(\|\chi_0 - \chi_2\|, k) + \gamma(\max_{h \geq 0} \|u_1(h) - u_2(h)\|)$$

(2)

II. LSTM NETWORKS

A. State space form

The LSTM network, with input $u \in \mathbb{R}^{n_u}$ and output $y \in \mathbb{R}^{n_y}$, is described by the following system of equations [12, 14].

$$x^+ = \sigma_g(W_f u + U_f \xi + b_f) \circ x +$$

$$+ \sigma_i(W_{ii} u + U_i \xi + b_i) \circ \sigma_i(W_e u + U_e \xi + b_e)$$

$$\xi^+ = \sigma_g(W_o u + U_o \xi + b_o) \circ \sigma_o(x^+)$$

$$y = \chi^T \xi + b_y$$

(3a) (3b) (3c)

The vector $\chi = [x^T \xi^T]^T$ is the state of the network, so that (3) can be rewritten in the general form (1). In the related terminology, $x \in \mathbb{R}^{n_x}$ is named hidden state, while $\xi \in \mathbb{R}^{n_x}$ is named output state (or cell).

In system (3), $\sigma_g(x) = \frac{1}{1 + e^{-x}}$ and $\sigma_o(x) = \tanh(x)$; when applied to a vector, we assume to apply them entry-wise. Also, $\circ$ is the element-wise (Hadamard) product. The terms $W_f, W_i, W_o, W_c \in \mathbb{R}^{n_x \times n_u}$, $U_f, U_i, U_o, U_c \in \mathbb{R}^{n_x \times n_x}, C \in \mathbb{R}^{n_x \times n_x}$ are weighting matrices and $b_f, b_i, b_o, b_c \in \mathbb{R}^{n_x}$, $b_y \in \mathbb{R}^{n_y}$ are biasing vectors.

In this work, we assume that the input is bounded, i.e., that

$$u \in \mathcal{U} = [-u_{\max}, u_{\max}]^{n_u}$$

(4)

Note that condition (4), if not imposed by physical input saturations, can be satisfied by means of a suitable normalization of the input variables.

B. Properties of the system functions and bounds on the variables

First of all, note that, in view of their definitions,

$$\sigma_g(t) \in (0, 1), \quad \forall t \in \mathbb{R}$$

$$\sigma_i(t) \in (-1, 1), \quad \forall t \in \mathbb{R}$$

(5a) (5b)

Also, $\sigma_g(t)$ and $\sigma_i(t)$ are Lipschitz functions [37] with Lipschitz constants $L_g = 0.25$ and $L_c = 1$, respectively, and they are both strictly monotonic. In view of (5), see (3),

$$\xi \in (-1, 1)^{n_x}, \text{ i.e. } \xi(j) \in (-1, 1), \forall j = 1, \ldots, n_x$$

(6)

Rewriting equation (3) for each entry of the state vectors we obtain:

$$x^+_j = \sigma_g(W_f u + U_f \xi + b_f)_j \circ x(j) +$$

$$+ \sigma_i(W_{ii} u + U_i \xi + b_i)_j \circ \sigma_i(W_e u + U_e \xi + b_e)_j$$

$$\xi^+_j = \sigma_g(W_o u + U_o \xi + b_o)_j \circ \sigma_o(x^+)_j$$

(7a) (7b)
Note that, in (7a), for each $j \in \{1, \ldots, n_x\}$
\[
|\sigma_g(W_j u + U_j \xi + b_j)(j)| 
\leq \|\sigma_g(W_j u + U_j \xi + b_j)\|_\infty 
\leq \max_{w \in W, \xi \in (-1)^n}\|\sigma_g(W_j u + U_j \xi + b_j)\|_\infty 
\leq \|\sigma_g(W_j u + U_j \xi + b_j)\|_\infty 
\leq \max_{w \in W, \xi \in (-1)^n}\|\sigma_g(W_j u + U_j \xi + b_j)\|_\infty 
\leq \|\sigma_g(W_j u + U_j \xi + b_j)\|_\infty 
\leq \sigma_g(\|W_j u_{\text{max}} + U_j b_j\|_\infty) = \bar{\sigma}_g^f 
\]  
(9)
where we relied on (4) and (6). With similar arguments we derive:
\[
|\sigma_g(W_i u + U_i \xi + b_i)(j)| \leq \bar{\sigma}_g^c = \sigma_g(\|W_i u_{\text{max}} + U_i b_i\|_\infty) 
\]  
(10)
\[
|\sigma_g(W_o u + U_o \xi + b_o)(j)| \leq \bar{\sigma}_g^o = \sigma_g(\|W_o u_{\text{max}} + U_o b_o\|_\infty) 
\]  
(11)
\[
|\sigma_c(W_i u + U_i \xi + b_i)(j)| \leq \bar{\sigma}_c^c = \sigma_c(\|W_i u_{\text{max}} + U_i b_i\|_\infty) 
\]  
(12)

Also, by analysing equation (7a), and recalling (8)–(12), we define an invariant set $\mathcal{X} \equiv \{x \in \mathbb{R} : |x| \leq \bar{\sigma}_g^c\}$ for $x_j$, i.e., such that
\[
|x_j(0)| \in \mathcal{X} \Rightarrow |x_j(t)| \in \mathcal{X}, \forall t \geq 0 
\]  
(13)
Thanks to this definition, we can bound $\sigma_c(x^+)(j)$ in (7b), namely
\[
|\sigma_c(x^+)(j)| \leq \sigma_c \left( \frac{\bar{\sigma}_g^c}{\max_{\bar{\sigma}_g^c}} \right) \leq \sigma_c \left( \frac{\bar{\sigma}_g^c}{\bar{\sigma}_g^c} \right) = \bar{\sigma}_c^c 
\]  
(14)

C. $\delta$ISS properties of LSTM networks

The following sufficient condition holds, ensuring $\delta$ISS of system (3).

**Theorem 1**: Denoting
\[
\alpha = \left[ \frac{1}{4} \|U_j\| \alpha \bar{\sigma}_g^c + \frac{1}{4} \|U_c\| \bar{\sigma}_o^o + \frac{1}{4} \|U_o\| \bar{\sigma}_o^c \right] 
\]  
the LSTM network in (3) is $\delta$ISS if $\rho(A) < 1$, where
\[
A = \begin{bmatrix} \bar{\sigma}_g^f & \alpha \bar{\sigma}_o^c + \frac{1}{4} \bar{\sigma}_c^c \|U_o\| \end{bmatrix} 
\]  
(15)

**Proof**: See the Appendix.

**Proposition 1**: The Schur stability of $A$ is ensured if the following inequalities hold:
\[
-1 + \left[ \alpha \bar{\sigma}_g^o \frac{1}{4} \bar{\sigma}_c^c \|U_o\| \right] < \frac{1}{4} \sigma_g^f \sigma_c^c \|U_o\| < 1 
\]  
(16)

**Proof**: See the Appendix

**Remark 1**: It is worth remarking that the conditions of Theorem 1 and Proposition 1 are explicit functions of the LSTM parameters. Therefore, they can be explicitly considered in the training phase of the network, so that $\delta$ISS of the estimated model can be forced.

### III. Control design

#### A. Observer design

The use of the LSTM network for model predictive control purposes calls for the availability of a state estimate of the plant, as represented in Figure 1. We propose the use of a tailored observer, guaranteeing a fast convergence of such estimate to the real state value.

The observer is a dynamical system with state
\[
\dot{\hat{x}} = [\bar{\sigma}_g^f \bar{\sigma}_c^c \bar{\sigma}_o^c] \dot{\hat{x}} + \hat{y} 
\]  
and output estimate $\hat{y}$, taking the following form:
\[
\hat{x}^+ = \sigma_g(W_j u + U_j \hat{\xi} + b_j + L_f(y - \hat{y})) \dot{\hat{x}} + 
\sigma_g(W_o u + U_o \hat{\xi} + b_o + L_o(y - \hat{y})) \dot{\hat{x}} + 
\sigma_c(W_i u + U_i \hat{\xi} + b_i + L_i(y - \hat{y})) \dot{\hat{x}} 
\]  
(17)
\[
\hat{\xi} = C \hat{x} + b_y 
\]  
where $L_f$, $L_o$ and $L_o$ are suitable observer gains to be properly selected.

**Theorem 2**: If the plant behaves according to $\tilde{f}$, and $\rho(A) < 1$ holds, then the observer (17) with gains $L_f$, $L_i$, and $L_o$ defined in such a way that
\[
\|U_f - L_f C\| \leq \|U_f\| \tag{18a} 
\]  
\[
\|U_i - L_i C\| \leq \|U_i\| \tag{18b} 
\]  
\[
\|U_o - L_o C\| \leq \|U_o\| \tag{18c} 
\]  
provides a state estimate converging to the real one, i.e., $\hat{x}(k) \to x(k)$ as $k \to \infty$.

**Proof**: See the Appendix.

Remark that, in order to fulfill (18), one can set $L_f = L_i = L_o = 0_{n_x}$, however, to ensure more reliable and possibly faster state estimation properties, one may solve the following optimization problems
\[
L_f = \arg \min_{L_f \in \mathbb{R}^{n_x}} \|U_f - L_f C\| \tag{19a} 
\]  
\[
L_i = \arg \min_{L_i \in \mathbb{R}^{n_x}} \|U_i - L_i C\| \tag{19b} 
\]  
\[
L_o = \arg \min_{L_o \in \mathbb{R}^{n_x}} \|U_o - L_o C\| \tag{19c} 
\]  

![Fig. 1. Control architecture](image-url)
The MPC scheme consists of solving, at each sampling time \( k \), the optimization problem

\[
\min_{U(k)} J(U(k))
\]

\[
s.t. u(k+i) \in \mathcal{U} \text{ for } i = 0, \ldots, N-1
\]

\( U(k) \) is the sequence, at step \( k \), of current and future \( N-1 \) control moves, i.e. \( U(k) = [u(k) \ldots u(k+N-1)]^T \) and \( N \) is the prediction horizon. The cost function reads:

\[
J(U(k)) = \sum_{i=0}^{N-1} (\|\chi(k+i) - \tilde{\chi}\|_Q + \|u(k+i) - \tilde{u}\|_R) + \|\Delta(k+N)\|_P^2
\]

where \( \chi = \rho(\Omega) \). At time step \( k \), the solution to the optimization problem is termed \( U(k) = [u(k)[k]^T \ldots u(k+N-1)[k]^T]^T \). Only its first element is applied to the system according to the Receding Horizon principle, i.e.,

\[
u(k) = u(k)[k]
\]

The following result holds, ensuring asymptotic stability of the equilibrium \( (\tilde{u}, \tilde{\chi}, \tilde{\chi}) \) under the proposed control law (24).

**Theorem 3**: If the plant behaves according to (3) with \( \rho(\Omega) < 1 \), the state observer is designed according to (17) and tuned fulfilling (18), then \( (\tilde{u}, \tilde{\chi}, \tilde{\chi}) \) is an asymptotically stable equilibrium under the control law (24).

**Proof**: See the Appendix.

**Remark 2**: Theorem 3 relies on the assumption that the plant is exactly described by the trained LSTM model (3), i.e. there is no model-plant mismatch. This assumption is rather standard in MPC results, is here supported by the fact that, if suitably trained, a NN has been proved to reproduce any plant dynamics (i.e. any function \( \varphi(\cdot, \cdot) \), see (1) ) with arbitrary precision [7, 36].

---

**B. Model Predictive Control design**

This section discusses the design of a predictive control scheme that takes advantage of the LSTM network (3) as a prediction model of the system. The objective of the controller is to stabilize the system toward a generic equilibrium point denoted by the triplet \( (\tilde{u}, \tilde{\chi}, \tilde{\chi}) \) (where, for consistency, also \( \tilde{u} \in \mathcal{U} \)) by suitably acting on the control input \( u \). With reference to a generic equilibrium \( (\tilde{u}, \tilde{\chi}, \tilde{\chi}) \), we define

\[
\Delta = \begin{bmatrix}
\|x - \tilde{x}\|_2 \\
\|\tilde{\chi} - \xi\|_2
\end{bmatrix} \in \mathbb{R}^2
\]

The MPC scheme consists of solving, at each sampling time \( k \), the solution to the optimization problem

\[
\min_{U(k)} J(U(k))
\]

\[
s.t. u(k+i) \in \mathcal{U} \text{ for } i = 0, \ldots, N-1
\]

\( U(k) \) is the sequence, at step \( k \), of current and future \( N-1 \) control moves, i.e. \( U(k) = [u(k) \ldots u(k+N-1)]^T \) and \( N \) is the prediction horizon. The cost function reads:

\[
J(U(k)) = \sum_{i=0}^{N-1} (\|\chi(k+i) - \tilde{\chi}\|_Q + \|u(k+i) - \tilde{u}\|_R) + \|\Delta(k+N)\|_P^2
\]

The terms \( \chi(k+i), i = 1, \ldots, N \) are the future state predictions, obtained by iterating (3) starting from the current state estimate \( \chi(k) = \tilde{\chi}(k) \). Matrices \( Q \geq 0 \) and \( R > 0 \) are tuning parameters. Matrix \( P > 0 \in \mathbb{R}^{2 \times 2} \) satisfies the Lyapunov condition - under the assumption that \( \rho(\Omega) < 1 \):

\[
A^T P A - P + \rho(f)I < 0
\]

where \( \rho = \rho(\Omega) \). At time step \( k \), the solution to the optimization problem is termed \( U(k) = [u(k)[k]^T \ldots u(k+N-1)[k]^T]^T \). Only its first element is applied to the system according to the Receding Horizon principle, i.e.,

\[
u(k) = u(k)[k]
\]

The following result holds, ensuring asymptotic stability of the equilibrium \( (\tilde{u}, \tilde{\chi}, \tilde{\chi}) \) under the proposed control law (24).

**Theorem 3**: If the plant behaves according to (3) with \( \rho(\Omega) < 1 \), the state observer is designed according to (17) and tuned fulfilling (18), then \( (\tilde{u}, \tilde{\chi}, \tilde{\chi}) \) is an asymptotically stable equilibrium under the control law (24).

**Proof**: See the Appendix.

**Remark 2**: Theorem 3 relies on the assumption that the plant is exactly described by the trained LSTM model (3), i.e. there is no model-plant mismatch. This assumption is rather standard in MPC results, is here supported by the fact that, if suitably trained, a NN has been proved to reproduce any plant dynamics (i.e. any function \( \varphi(\cdot, \cdot) \), see (1) ) with arbitrary precision [7, 36].

---

**IV. ILLUSTRATIVE EXAMPLE**

The benchmark example here considered to test the described identification and control algorithm is a PH neutralization process [18], composed of two tanks, namely Tank 1 and Tank 2, see also Figure 2.

Tank 2 is fed by an acid stream \( q_1 \) and outputs a flow \( q_{1e} \), but this hydraulic dynamics is neglected being much faster than the others involved, so that \( q_1 = q_{1e} \). Tank 1, also called reactor tank, is instead fed by three flows, namely \( q_3 \), a buffer flow \( q_2 \) and an alkaline flow \( q_4 \). \( q_1 \) and \( q_4 \) are not manipulated variables, and represent disturbances, whereas a controlled valve modulates \( q_3 \). On the output side is flow \( q_4 \), where the PH is measured. The objective of the control scheme is to stabilize the PH concentration to a desired value.

The plant is characterized by the following set of differential equations with a constraint [18]:

\[
x(t) = f_1(x(t)) + f_2(x(t))u(t) + f_3(x(t))d(t) = 0
\]

where

\[
f_1(x(t)) = \left[ \frac{q_1}{A_{x_3}}(W_{d_1} - x_1) , \frac{q_1}{A_{x_3}}(W_{d_2} - x_2) , \frac{1}{A_1}(q_1 - C_d(x_3 + z)) \right]^T
\]

\[
f_2(x(t)) = \left[ \frac{1}{A_{x_3}}(W_{d_1} - x_1) , \frac{1}{A_{x_3}}(W_{d_2} - x_2) , \frac{1}{A_1} \right]^T
\]

\[
f_3(x(t)) = \left[ \frac{1}{A_{x_3}}(W_{d_2} - x_2) , \frac{1}{A_{x_3}}(W_{d_2} - x_2) , \frac{1}{A_1} \right]^T
\]

and

\[
c(x,y) = x_1 + 10^{-14} + 10^{-7} + x_2 - 1 + 2 \cdot 10^{-7} \cdot p K_2
\]

\( p K_1 \) and \( p K_2 \) are the first and second dissociation constants of the weak acid \( H_2CO_3 \). The nominal values of the model parameters are given in Table 4, where \( [M] = [mol/L] \).

Overall, the simplified model considered is of order three, with one input and one output.

**A. Identification**

The simulator of the plant is forced with a multilevel pseudo-random signal (MPRS), so as to properly excite
the system, and the input-output data are recorded with a sampling time $T_s = 10s$ so as to collect about 30-40 samples in a step response. Also, to test the algorithm in a more realistic scenario, a measurement white noise is added to the output variable, with power $3 \times 10^{-4}W$. The available dataset consists then in 1000 $(u, y)$ samples for the training set, and as many for the validation dataset, where the modeling performances are evaluated as common practice, see also [32].

An LSTM model [5] with $n_x = 20$ neurons is trained, and conditions (16) are enforced as constraints on the internal weights in the training optimization problem, thus ensuring $\delta$ISS.

The modeling performances over the validation dataset are reported in Figure 3 where a simulation of the trained network is shown, initialized from a random value, and forced by the input $u$. A quantitative performance index is the FIT [%] value, computed as

$$FIT = 100 \left(1 - \frac{\|Y - \bar{Y}_{\text{LSTM}}\|}{\|Y - \bar{Y}\|}\right)$$

where $Y$ collects the output samples of the dataset, $\bar{Y}$ is its average and $\bar{Y}_{\text{LSTM}}$ collects the output simulation of the trained LSTM. Over the validation dataset, the FIT scores 97%, thus confirming remarkable modeling properties.

**B. Control**

The designed observer follows (17) and is tuned according to (19), thus guaranteeing a realizable state estimate to the MPC controller.

The testing experiment is a reference tracking one. More specifically, the controller is started at time 500s and it required to track a setpoint reference $\bar{y} \in \{7, 8, 7.5, 6.5, 7\}$ and stabilize the associated equilibrium. Exploiting the designed observer, and assuming to know the pairs $(\bar{u}, \bar{y})$, the corresponding steady state $\bar{z}$ is obtained by simulation, in view of the guaranteed convergence properties of the observer itself. The weighting matrices in the cost function of the controller are $Q = 0.1I_{2n_y}$, $R = 0.1$, and $P$ is computed according to [23].

The closed-loop trajectory is reported in Figure 4 which shows that the controller is able to effectively manage the plant, fulfilling control constraints and improving the transient responses. In particular, note that around 1800s the input is saturated to its upper bound. To confirm the validity of the estimate provided by the observer, in Figure 5 the output estimate $\hat{y}$ and the system output are reported, showing convergence of the estimate, save for a static mismatch due to the model (LSTM) - plant (pH simulator) gain mismatch when pH $\approx 8$.

### V. Conclusion

In this paper Long Short Term Memory networks have been investigated from a system theoretical perspective, and sufficient conditions for their $\delta$ISS stability have been provided in terms of their internal weights. A novel formulation of the optimization problem to train the NN, including constraints, has been employed. The obtained NN has then been used as a prediction model in a MPC scheme endowed with an observer to suitably provide the initial state estimate, with guaranteed convergence of the estimate and asymptotic stability of the closed-loop equilibrium. Numerical result on a nonlinear SISO benchmark confirm the theoretical findings in the case of a tracking problem.

Future work is concerned with the extension of the analysis to other classes of LSTM, e.g. the peephole one.
including feedback of $x$ in the state equations, and with the development of robust control algorithms explicitly taking into account the model-plant mismatch.

VI. APPENDIX

Proof of Theorem 1 To prove the theorem, we will use the following properties.

Property 1: Given vectors $v_1, v_2 \in \mathbb{R}^n$, $v_1 \circ v_2 = \text{diag}(v_1)v_2$.

Property 2: Given a diagonal matrix $A$, $\|A\| = \rho(A)$, and the eigenvalues of $A$ is its diagonal entries.

Now we compute the evolution of the upper bound of the norms of the two components of the state $\chi$, namely $x$ and $\xi$. We proceed by addressing the two subvectors separately.

- $x^+_1 - x^+_2 =$

$$= \sigma_g(W_fu_1 + U_f\xi_1 + b_f) \circ x_1$$
$$+ \sigma_g(W_fu_1 + U_f\xi_1 + b_i) \circ \sigma_e(W_cu_1 + U_c\xi_1 + b_c)$$
$$- \sigma_g(W_fu_2 + U_f\xi_2 + b_f) \circ x_2$$
$$+ \sigma_g(W_fu_2 + U_f\xi_2 + b_i) \circ \sigma_e(W_cu_2 + U_c\xi_2 + b_c)$$

$$= \sigma_g(W_fu_1 + U_f\xi_1 + b_f) \circ (x_1 - x_2)$$
$$+ x_2 \circ [\sigma_g(W_fu_1 + U_f\xi_1 + b_f) - \sigma_c(W_fu_2 + U_f\xi_2 + b_f)]$$
$$+ \sigma_g(W_fu_1 + U_f\xi_1 + b_i) \circ [\sigma_e(W_cu_1 + U_c\xi_1 + b_c)$$
$$- \sigma_e(W_cu_2 + U_c\xi_2 + b_c)]$$
$$+ \sigma_e(W_cu_2 + U_c\xi_2 + b_c) \circ [\sigma_g(W_fu_1 + U_f\xi_1 + b_i) - \sigma_c(W_fu_2 + U_f\xi_2 + b_i)]$$

Recalling the upper bounds (8)-(12), Lipschitzianity of $\sigma_c(\cdot)$ and $\sigma_g(\cdot)$ and taking the norms both sides, we write, in view of Properties [1] and [2] $\|x^+_1 - x^+_2\| \leq$

$$\leq \sigma_g \|x_1 - x_2\| + \frac{\sigma_g}{1 - \sigma_g} \frac{1}{4} \left( \||W_f\||u_1 - u_2\| + \||U_f\||\|\bar{\xi}_1 - \bar{\xi}_2\| \right)$$
$$+ \frac{\sigma_c}{1 - \sigma_c} \frac{1}{4} \left( \||W_c\||u_1 - u_2\| + \||U_c\||\|\bar{\xi}_1 - \bar{\xi}_2\| \right)$$
$$\leq \sigma_g \|x_1 - x_2\| + \alpha \|\bar{\xi}_1 - \bar{\xi}_2\| + \beta \|u_1 - u_2\|$$

(28)

where $\alpha = \frac{1}{2} \||W_f\|\|\frac{\sigma_g}{1 - \sigma_g} + \frac{\sigma_c}{1 - \sigma_c}\||W_c\| + \frac{1}{4} \||U_f\|\|\bar{\xi}_1 - \bar{\xi}_2\|)$
$$\text{and } \beta = \frac{1}{2} \||W_f\|\|\frac{\sigma_g}{1 - \sigma_g} + \frac{\sigma_c}{1 - \sigma_c}\||W_c\| + \frac{1}{4} \||W_f\||\bar{\xi}_1 - \bar{\xi}_2\|)$

(29)

Also $\bar{\xi}_1 - \bar{\xi}_2 =$

$$= \sigma_g(W_fu_1 + U_f\xi_1 + b_o) \circ \sigma_c(x_1) - \sigma_g(W_fu_2 + U_f\xi_2 + b_o) \circ \sigma_c(x_2)$$
$$= \sigma_g(W_fu_1 + U_f\xi_1 + b_o) \circ \sigma_e(x_1) - \sigma_c(x_2) + \sigma_e(x_1) \circ [\sigma_g(W_fu_1 + U_f\xi_1 + b_o) - \sigma_g(W_fu_2 + U_f\xi_2 + b_o)]$$

(30)

By recalling the bounds (8)-(12), taking the norm both sides, we write $\|\bar{\xi}_1 - \bar{\xi}_2\| \leq$

$$\leq \sigma_g \|x_1 - x_2\| + \frac{1}{4} \left( \||W_f\||u_1 - u_2\| + \||U_o\||\|\bar{\xi}_1 - \bar{\xi}_2\| \right)$$
$$\leq \sigma_g \left( \|x_1 - x_2\| + \alpha \|\bar{\xi}_1 - \bar{\xi}_2\| + \beta \|u_1 - u_2\| \right)$$
$$+ \frac{1}{4} \left( \||W_f\||u_1 - u_2\| + \||U_o\||\|\bar{\xi}_1 - \bar{\xi}_2\| \right)$$
$$\leq \sigma_g \left( \|x_1 - x_2\| + \left[ \alpha \sigma_g + \frac{1}{4} \sigma_c \||U_o\||\right] \|\bar{\xi}_1 - \bar{\xi}_2\| + \right.$$

$$\left. \beta \sigma_g + \frac{1}{4} \sigma_c \||W_f\|| \||u_1 - u_2\| \right)$$

(31)

Grouping inequalities (28) and (30), we obtain that

$$\left\| \frac{\|x^+_1 - x^+_2\|}{\|\bar{\xi}_1 - \bar{\xi}_2\|} \right\| \leq A \left( \|x_1 - x_2\| \right) + B \|u_1 - u_2\|$$

(32)

where $A$ is defined in (15) and $B = \left[ \beta \sigma_g + \frac{1}{4} \sigma_c \||W_f\|| \right]$.

We now finally show that the stability of $A$ implies the $\delta$ISS property of system (3). Let us first recall that

$$\||x_1 - x_2\| \leq \||x_1 - x_2\| \leq \||\bar{\xi}_1 - \bar{\xi}_2\| \leq \||\bar{\xi}_1 - \bar{\xi}_2\|$$

Consider system (31), we obtain, by iteration, that

$$\left\| \frac{\|x_1(k) - x_2(k)\|}{\|\bar{\xi}_1 - \bar{\xi}_2(k)\|} \right\| \leq A^k \left( \|x_0 - x_0\| \right) + \sum_{i=0}^{k-1} A^{k-i-1} B \|u_1(i) - u_2(i)\|$$

(33)

Fig. 5. Comparison of real (black) and estimated output (grey)
By taking the norm both sides we compute
\[
\|\chi_1(k) - \chi_2(k)\| \leq \|A^k\| \|\chi_{01} - \chi_{02}\| + \\
\|\sum_{i=0}^{k-1} A^{k-i-1} B\| \|u_1(i) - u_2(i)\| \tag{33}
\]

With standard norm arguments, and recalling that \(A\) is Schur stable, there exist constants \(\mu \geq 1\) and \(\lambda \in (0, 1)\) such that
\[
\|\chi_1(k) - \chi_2(k)\| \leq \mu \lambda^k \|\chi_{01} - \chi_{02}\| + \\
\|\sum_{h=0}^{k} (I_2 - A)^{-1} B\| \max_{h \geq 0} \|u_1(h) - u_2(h)\| \tag{34}
\]
i.e., \(\chi\) is \(\delta\)ISS, according to Definition 2.

Proof of Proposition [1]

To characterize the stability of \(A\), defined in (15), we can compute its characteristic equation, i.e.,
\[
p(\lambda) = \det(\lambda I_2 - A) = \lambda^2 + a\lambda + b = 0 \tag{35}
\]
where \(a = -\sigma_g^c - \alpha \sigma_g^o - \frac{1}{4} \sigma_y^o \|U_o\|\) and \(b = \frac{1}{4} \sigma^o \sigma^c \|U_o\|\). We rely on Jury’s criterion [27], providing a necessary and sufficient condition, to enforce stability of \(A\). The Jury table of \(p(\lambda)\) is
\[
\begin{array}{c|cc|c}
& a & b & (1-b)^2 \\
\hline
a & 1 & -b^2 & \frac{1-b^2}{(1-b)^2} \\
b & a(1-b) & b \end{array}
\tag{36}
\]

Jury’s criterion requires to force the first column to have all positive entries, that leads, with standard arguments and recalling that \(a < 0\), to the set of conditions:
\[
\begin{cases}
b^2 < 1 \\
1 - b^2 > -a(1-b)
\end{cases} \tag{37}
\]
which can be further synthesized in
\[
-1 - a < b < 1 \tag{38}
\]
Finally, by properly replacing \(a\) and \(b\) in (38), condition (16) is obtained.

Proof of Theorem 2

Let us define the error variables \(e_x = x - \hat{x}\), \(e_\xi = \xi - \hat{\xi}\) and compute their evolution over time. In particular:
\[
e^x_+ = x^+ - \hat{x}^+ = \\
= \sigma_g(W_f u + U_f \xi + b_f) \circ x + \\
+ \sigma_g(W_d u + U_d \xi + b_d) \circ \sigma_c(W_c u + U_c \xi + b_c) - \\
\left\{ \sigma_g(W_f u + U_f \xi + b_f + L_f(y - \hat{y})) \circ \hat{x} + \\
+ \sigma_g(W_d u + U_d \xi + b_d + L_d(y - \hat{y})) \circ \sigma_c(W_c u + U_c \xi + b_c) \right\} \tag{39}
\]
Computing the norm of both sides we obtain \(\|e^x_+\| \leq \|
\]
\[
\frac{\bar{\sigma}^o \|e_x\| + \bar{\sigma}^c \|e_\xi\|}{1 - \bar{\sigma}^c} \] + \frac{1}{4} \|U_f - L_f C\| \|e^x_+\| + \bar{\sigma}^c \|U_c\| \|e_\xi\| \\
= \frac{1}{4} \|U_f - L_f C\| \|e_x\| \tag{40}
\]
which can be obtained thanks to (38).

\[
e^\xi_+ = \xi^+ - \hat{\xi}^+ = \\
= \sigma_g(W_o u + U_o \xi + b_o) \circ \sigma_c(x^+) - \\
\sigma_g(W_o u + U_o \xi + b_o + L_o(y - \hat{y})) \circ \sigma_c(x) \tag{41}
\]
Computing the norm of both sides we obtain \(\|e^\xi_+\| \leq \|
\]
\[
\frac{\bar{\sigma}^o \|e_x\| + \bar{\sigma}^c \|e_\xi\|}{1 - \bar{\sigma}^c} \] + \frac{1}{4} \|U_o - L_o C\| \|e^\xi_+\| + \bar{\sigma}^o \|U_o\| \|e^o\| \\
= \frac{1}{4} \|U_o - L_o C\| \|e_x\| \tag{42}
\]
which, again, is obtained thanks to (18).

We can write, from (40) and (42)
\[
\begin{bmatrix} \|e^x_+\| \\ \|e^\xi_+\| \end{bmatrix} \leq A \begin{bmatrix} \|e_x\| \\ \|e_\xi\| \end{bmatrix} \tag{43}
\]
Therefore, the Schur stability of \(A\) guarantees the exponential convergence of the estimation error.

Proof of Theorem 3

The following property will be used in the sequel.

Property 3: Given two vectors \(a, b \in \mathbb{R}^n\) and a positive definite matrix \(M > 0\), it holds that
\[
\|a + b\|^2_M \leq (1 + v^2)\|a\|^2_M + (1 + \frac{1}{v^2})\|b\|^2_M
\]
for any \(v \neq 0\).
Assume that, at time step $k$, the optimal solution $U(k|k)$ to the MPC problem is obtained, and the input value $u(k) = u(k|k)$ is applied to the system. We denote with $\chi(k+i|k)$, $i = 0, \ldots, N$ the state trajectory obtained iterating (3) with initial condition $\chi(k|k) = \tilde{\chi}(k)$, and using $U(k|k)$ as input sequence.

Similarly, we denote $\Delta(k+N|k) = \left[ \|x(k+N[k] - \tilde{x})\|_q + \|\Delta(k+N|k)\|_p \right]^2$.

We will consider the optimal value of $J$ (denoted $J^*(k)$) as a candidate Lyapunov function to analyze the stability properties of the MPC control algorithm. We compute that

$$J^*(k) = \sum_{i=0}^{N-1} \left( ||\chi(k+i|k) - \tilde{\chi}||_q^2 + ||\Delta(k+N|k)\|_p^2 \right).$$

Note first that

$$J^*(k) \geq \gamma_1 \|\tilde{\chi}(k) - \tilde{\chi}\|^2$$

(44)

where $\gamma_1 = \lambda_\text{max}(Q)$. Secondly, note that $u(k+i) = \bar{u}$ is a possibly suboptimal but feasible control input for all $i = 0, \ldots, N - 1$. We denote with $\chi^o(k+i|k) = [\chi^o(k+i|k) \bar{u}]^T$, $i = 0, \ldots, N$ the state trajectory obtained iterating (3) with initial condition $\chi(k|k) = \tilde{\chi}(k)$, and using $\{u(k) = \bar{u}, \ldots, u(k+N-1) = \bar{u}\}$ as input sequence. We obtain that

$$J^*(k) \leq \sum_{i=0}^{N-1} ||\chi^o(k+i|k) - \tilde{\chi}||_q^2 + ||\Delta'(k+N|k)\|_p^2$$

where $\Delta'(k+N|k) = [||x^o(k+N[k] - \tilde{x})||_q + ||\bar{x}||_q]$. Further note that $||\Delta'(k+N|k)\|_p^2 \leq \lambda_\text{max}(P)||\Delta'(k+N|k)\|_q^2 = \lambda_\text{max}(P)||\chi^o(k+N[k] - \tilde{\chi}||_q^2$. Secondly, similarly to the proof of Theorem 4, since $\rho(A) < 1$ there exist $\mu \geq 0$ and $\lambda \in (0, 1)$ such that, for all $i \geq 0$

$$||\chi^o(k+i|k) - \tilde{\chi}|| \leq \mu \lambda^i ||\tilde{\chi}(k) - \tilde{\chi}||$$

This implies that there exists a constant $\gamma_2$ such that

$$J^*(k) \leq \gamma_2 ||\tilde{\chi}(k) - \tilde{\chi}\|^2.$$

At time $k = 1$ (with some abuse of notation, but for the sake of simplicity), we denote with $\chi(k+i+1)$, $i = 1, \ldots, N+1$ the possibly suboptimal state trajectory obtained iterating (3) with initial condition $\chi(k+1|k+1) = \tilde{\chi}(k+1)$, and using $\{u(k+1|k), \ldots, u(k+N-1|k), \bar{u}\}$ as input sequence. Note that $\chi(k+1|k+1) = \tilde{\chi}(k+1) \neq \phi(\tilde{\chi}(k), u(k|k)) = \chi(k+1|k)$, We denote $\epsilon(k+1|k+1) = \tilde{\chi}(k+1) - \chi(k+1|k)$ and $\epsilon_\Delta(k+1|k+1) = \Delta(k+1|k+1) - \Delta(k+1|k)$, Similarly, we define $\epsilon(k+i+1|k+1) = \chi(k+i+1|k+1) - \chi(k+1|k)$ and

$$\epsilon_\Delta(k+i+1|k+1) = \Delta(k+i+1|k+1) - \Delta(k+i+1|k)$$

for all $i = 0, \ldots, N - 1$. The optimal value $J^*(k+1)$ satisfies the following inequality.

$$J^*(k+1) \leq \sum_{i=0}^{N-1} \left( ||\chi(k+i+1|k+1) - \tilde{\chi}||_q^2 + ||u(k+i+1|k+1) - \bar{u}||_q^2 \right) + ||\Delta(k+N+1|k+1)\|_p^2$$

$$\leq \sum_{i=0}^{N-1} \left( ||\chi(k+i+1|k) - \tilde{\chi}||_q^2 + \epsilon(k+i+1|k+1)\|_q^2 + \epsilon_\Delta(k+N+1|k+1)\|_p^2 \right)$$

Therefore $J^*(k+1) - J^*(k) \leq \ldots$

$$- ||\chi(k) - \tilde{\chi}||_q^2 - ||u(k) - \bar{u}||_q^2$$

$$+ \sum_{i=0}^{N-1} \left( ||\chi(k+i|k) - \tilde{\chi} + \epsilon(k+i|k+1)\|_q^2 - ||\chi(k+i|k) - \tilde{\chi}\|_q^2 \right)$$

$$- ||\Delta(k+N|k)\|_p^2 + ||\chi(k+N|k) - \tilde{\chi} + \epsilon(k+N|k+1)\|_q^2$$

$$+ ||u(k+N|k+1) - \bar{u}||_q^2$$

$$+ ||\Delta(k+N+1|k) + \epsilon_\Delta(k+N+1|k+1)\|_p^2 \right)$$

(45)

We now consider the different additive terms at the right hand side of inequality (45). First, we write

$$\sum_{i=0}^{N-1} \left( ||\chi(k+i|k) - \tilde{\chi} + \epsilon(k+i|k+1)\|_q^2 - ||\chi(k+i|k) - \tilde{\chi}\|_q^2 \right)$$

$$+ \sum_{i=0}^{N-1} \left( ||\chi(k+i|k) - \tilde{\chi}||_q^2 - ||\chi(k+i|k) - \tilde{\chi} + \epsilon(k+i|k+1)\|_q^2 \right)$$

$$+ 2(\epsilon(k+i|k) - \tilde{\chi})^T Q \epsilon(k+i|k+1)$$

(46)

Also, we compute that, in view of Property 3

$$- ||\Delta(k+N|k)\|_p^2 + ||\chi(k+N|k) - \tilde{\chi} + \epsilon(k+N|k+1)\|_q^2$$

$$+ ||\Delta(k+N+1|k) + \epsilon(k+N+1|k+1)\|_p^2$$

$$\leq - ||\Delta(k+N|k)\|_p^2 + (1 + \rho^2 \sum_{i=0}^{N-1} \left( ||\chi(k+N|k) - \tilde{\chi}\|_q^2 + \frac{1}{1 + \rho^2} ||\epsilon(k+N|k+1)\|_q^2 \right)$$

$$+ \frac{1}{1 + \rho^2} \epsilon_\Delta(k+N+1|k+1)\|_p^2$$

$$+ \left( 1 + \frac{1}{\rho^2} \right) \|\epsilon(k+N|k+1)\|_q^2 + \|\epsilon_\Delta(k+N+1|k+1)\|_p^2 \right)$$

(47)

Now note that

$$\|\Delta\| = \left[ \left\| x - \tilde{x} \right\|_q^2 + \left\| \tilde{x} - \tilde{x} \right\|_q^2 \right]^2 = \left\| \chi - \tilde{\chi} \right\|_q^2$$

(46)

Also, in view of the $\delta$ISS property of the system, it holds that

$$\Delta(k+N+1|k) \leq A \Delta(k+N|k)$$

(47)
Therefore we write

\[-\|\Delta(k+N|k)\|_{\hat{P}}^2 + (1+\rho^2)(\|\chi(k+N|k) - \chi\|_{\hat{P}}^2
+ \|\Delta(k+N+1|k)\|_{\hat{P}}^2) \leq - \|\Delta(k+N|k)\|_{\hat{P}}^2 + \|\Delta(k+N+1|k)\|_{\hat{P}}^2 \leq \|\Delta(k+N|k)\|_{\hat{P}}^2 \]

By construction $A^TPA - P < -qI_2$, see (23), then we can always select a value of $\rho > 0$ small enough to obtain $A^TPA - P + qI_2(1+\rho^2) < 0$. Overall, we obtain that

\[J^*(k+1) - J^*(k) \leq -\|\chi(k) - \chi\|_{\hat{P}}^2 - \|u(k) - \bar{u}\|_{\hat{K}}^2 + (a), \]

where $a = \sum_{i=1}^{N-1} \left(\|\varepsilon(i+k|k+1)\|_{\hat{Q}}^2 + 2(\chi(i+k|k) - \chi)^TQ(e(i+k|k+1))\right) + \left(1 + \frac{1}{\rho^2}\right)\left(\|\varepsilon(k+N|k+1)\|_{\hat{P}}^2 + \|\varepsilon_a(k+N+1|k+1)\|_{\hat{P}}^2\right)$. Overall, we can, for simplicity, define an upper bound to (a) as

\[a \leq \alpha_1 \sum_{i=1}^{N-1} \|\varepsilon(i+k|k+1)\|_{\hat{Q}}^2 + \alpha_2 \sum_{i=1}^{N-1} \|\varepsilon(i+k|k+1)\|_{\hat{P}}^2 \]

This implies that

\[\|\varepsilon(k+N|k+1)\|_{\hat{P}} \leq \alpha_3 \|\chi(k) - \chi\|_{\hat{P}}^2 \]

By combining inequalities (48) and (53) we eventually obtain that there exists a $A^TQ$ function $\hat{P}$ and a constant $\gamma_2 > 0$ such that $J^*(k+1) - J^*(k) \leq -\gamma_2 \|\chi(k) - \chi\|_{\hat{P}}^2 + \tilde{\beta}(\|\varepsilon(0)\|, k)$, where $\tilde{\beta}(\|\varepsilon(0)\|, k)$ is exponentially decreasing with respect to its second argument $k$. Along the same lines of reasoning of Theorem 3 in [34], we can prove asymptotic stability of the equilibrium point denoted by the triplet $(\bar{u}, \tilde{\chi}, \tilde{y})$.

\[\square\]

**REFERENCES**

[1] Massinissa Amrouche, Deka Shankar Anand, Aleksandra Lekić, Vicenc Rubies Royo, Elana Teresa Chai, Dušan M. Stipanović, Boris Murmann, and Claire J Tomlin. Long short-term memory neural network equilibria computation and analysis. 2018.

[2] L. B. Armenio, E. Terzi, M. Farina, and R. Scattolini. Model predictive control design for dynamical systems learned by echo state networks. IEEE Control Systems Letters, 3(4):1044–1049, Oct 2019.

[3] Anil Aswani, Huminer Gonzalez, S. Shankar Santra, and Claire Tomlin. Provably safe and robust learning-based model predictive control. Automatica, 49(5):1216–1226, 2013.

[4] T. Bonetti and D. Saccani. Modeling and learning-based hybrid predictive control with recurrent neural networks of the cooling station of a large business center. Technical report, MS thesis Politecnico di Milano, 2019. Supervisor: R. Scattolini, Co-Supervisor: E. Terzi.

[5] Douglas A Bristow, Marina Tharayil, and Andrew G Alleyne. A survey of iterative learning control. IEEE control systems magazine, 26(3):96–114, 2006.

[6] M.C. Campi, A. Lecchini, and S.M. Savaresi. Virtual reference feedback tuning: a direct method for the design of feedback controllers. Automatica, 38(8):1337 – 1346, 2002.

[7] Tianping Chen and Hong Chen. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems. IEEE Transactions on Neural Networks, 6(4):911–917, 1995.

[8] S. A. Deka, D. M. Stipanović, B. Murmann, and C. J. Tomlin. Long-term memory neural network stability and stabilization using linear matrix inequalities. In 2019 IEEE International Symposium on Circuits and Systems (ISCAS), pages 1–4, May 2019.

[9] Shankar A. Deka, Dušan M. Stipanović, Boris Murmann, and Claire J. Tomlin. Global asymptotic stability and stabilization of long short-term memory neural networks with constant weights and biases. Journal of Optimization Theory and Applications, 181(1):231–243, Apr 2019.

[10] A Delgado, C Kambhampati, and Kevin Warwick. Dynamic recurrent neural network for system identification and control. IEEE Proceedings-Control Theory and Applications, 142(4):307–314, 1995.

[11] Bayer F., M. Burger, and F. Allgower. Discrete-time incremental ISS: A framework for robust NMPC. In Control Conference (ECC), 2013 European, pages 2068–2073. IEEE, 2013.

[12] Felix A Gers and E Schmidhuber. Lstm recurrent networks learn simple context-free and context-sensitive languages. IEEE Transactions on Neural Networks, 12(6):1333–1340, 2001.

[13] Felix A Gers, Jürgen Schmidhuber, and Fred Cummins. Learning to forget: Continual prediction with lstm. In Advances in neural information processing systems, pages 545–552. 2002.

[14] Alex Graves, Abdel-rahman Mohamed, and Geoffrey Hinton. Speech recognition with deep recurrent neural networks. In 2013 IEEE international conference on acoustics, speech and signal processing, pages 6645–6649. IEEE, 2013.

[15] Alex Graves and Jürgen Schmidhuber. Offline handwriting recognition with multidimensional recurrent networks. In Advances in neural information processing systems, pages 545–552. 2009.

[16] Klaus Greff, Rupesh K Srivastava, Jan Koutník, Bas R Steunebrink, and Jürgen Schmidhuber. Lstm: A search space odyssey. IEEE transactions on neural networks and learning systems, 28(10):2222–2232, 2016.
