

Estimation and MPC control based on gated recurrent unit neural networks with unknown disturbances

Eva Masero¹, Fabio Bonassi², Alessio La Bella¹, Riccardo Scattolini¹

Abstract—This paper proposes a nonlinear model predictive control (NMPC) approach for incrementally input-to-state stable gated recurrent units (GRU) neural networks affected by state and output disturbances. In particular, a Luenberger-like observer is designed for state and disturbance estimation with guaranteed convergence properties. This paves the way for the design of an NMPC regulator capable of rejecting unknown piecewise-constant disturbances. The method is tested in simulation on a nonlinear benchmark system, *i.e.*, a chemical reaction process, showing promising results.

Index Terms—Nonlinear Systems; Recurrent Neural Networks; Model Predictive Control; Stability

I. INTRODUCTION

Model Predictive Control (MPC) is a well-established control strategy that stands out for its capacity to manage input, state, and output constraints, as well as the conceptually straightforward extension to nonlinear system models. A key ingredient to achieve a well-performing MPC regulator is the accuracy of the dynamical model of the system under control, not always available in practice. Hence, a system identification procedure is often needed to retrieve a reliable dynamical model of the system from data, resulting in the so-called learning-based MPC [1].

Among the many structures available to learn black-box models for MPC design, researchers have recently considered Recurrent Neural Network (RNN) as Gated Recurrent Units (GRU) and Long Short-Term Memory Networks (LSTM) [2], as they can approximate any arbitrary dynamical system while avoiding gradient problems affecting other RNN models [3].

This choice has been shown to effectively combine the modeling capabilities of RNNs with the flexibility of Nonlinear MPC (NMPC), leading to control schemes with suitable performance in many applications [4], [5]. Moreover, theoretical frameworks based on the idea of training RNN models with Incremental Input-to-State Stability (δ ISS) have been proposed in [6], allowing the design of NMPC strategies with nominal closed-loop stability guarantees [7], [8].

A limitation of these stabilizing MPC strategies is their static performances, *i.e.*, their accuracy in regulating the system's output to constant setpoints. This accuracy might indeed be limited by the possible presence of plant-model mismatch and unknown disturbances affecting the plant, which calls for the design of offset-free NMPC strategies.

* Financial support by the Next-Generation EU (Italian PNRR-M4 C2, Invest 1.3-D.D.1551.11-10-2022, PE00000004) under Grant CUP MICS D43C22003120001 is gratefully acknowledged.

¹ E. Masero, A. La Bella and R. Scattolini are with the Dept. of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy.

² F. Bonassi is with the Dept. of Information Technology, Uppsala University, Uppsala, Sweden.

In the MPC literature, several strategies have been proposed to achieve offset-free MPC control [9]. One approach relies on the internal model principle [10], in which an integral action on the output tracking error is introduced and a stabilizing control law is designed for the augmented system given by the model plus the integrator. A second approach involves augmenting the system model with disturbance dynamics [11]. This augmented model enables the design of a state observer that estimates both the system state and the disturbance, which can be then used to design an NMPC law that compensates for the disturbance. Crucial to the success of this approach is the model adopted for the disturbance and the interaction of the disturbance with the system [12], as well as the design of an observer capable of actually estimating the disturbance. Adopting a disturbance model that only affects the system output transformation, [13] recently proposed a robust disturbance estimation-based offset-free NMPC strategy for GRU models.

While focusing solely on output disturbances can achieve satisfactory control performance, it may overlook state disturbances, which often affect real systems. Suitably estimating state disturbances is crucial not only for effective control but also for designing monitoring and fault detection algorithms. Moreover, learning the effect of disturbance on states allows one to exploit available knowledge about the physical system, aligning with the principles of physics-informed machine learning, which has been shown to yield improved models and faster learning procedures.

In this paper, we propose an offset-free NMPC strategy based on GRU models in which the disturbance affects both the states and the outputs. Specifically, we assume that disturbances are not measurable in closed-loop online operations. A GRU model that is δ ISS with respect to inputs and disturbances is proposed and then augmented with the disturbance dynamics. We present an observer with convergence guarantees for the augmented system, enabling the reconstruction of the model's state and the disturbance. Finally, a stabilizing MPC law is adopted for the extended system, which leads to an offset-free control architecture.

Notation: For vector $v \in \mathbb{R}^n$, $v_{(j)}$ denotes its j -th component, v^\top its transpose, and $\|v\|_p$ its ℓ_p -norm. The element-wise product between vectors u and v is $u \circ v$. For time-dependent vectors, we omit the time index k when there is no ambiguity, *i.e.*, $v = v_k$ and $v^+ = v_{k+1}$. We denote sequences by $v_{k_1:k_n} = \{v_{k_1}, \dots, v_{k_n}\}$. The hyperbolic tangent and sigmoid activation functions are denoted as $\phi(v) = \tanh(v)$ and $\sigma(v) = \frac{1}{1+e^{-v}}$, respectively. For

matrix A , $\|A\|_p$ is its induced p -norm and $\rho(A)$ its spectral radius. I denotes the identity matrix of proper dimensions. A continuous function $\alpha : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is a class \mathcal{K} if $\alpha(s) > 0$ for all $s > 0$ and $\alpha(0) = 0$, and $\alpha \in \mathcal{K}$ is also a \mathcal{K}_∞ function if $\alpha(s) \rightarrow \infty$ for $s \rightarrow \infty$. A continuous function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{Z}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is a class \mathcal{KL} if $\beta(s, k) \in \mathcal{K}$ is strictly decreasing in k for $s < 0$, and $\beta(s, k) \rightarrow 0$ as $k \rightarrow \infty$ for $s > 0$.

II. GRU MODEL WITH DISTURBANCES

Let us consider a single-layer GRU model described in discrete-time state-space form:

$$\Sigma(\Phi) : \begin{cases} x^+ = f_\Sigma(x, u, d; \Phi) = z \circ x + (1 - z) \circ r \\ y = g_\Sigma(x, d; \Phi) = V_o d + U_o x + b_o \end{cases} \quad (1a)$$

where $x \in \mathbb{R}^{n_x}$ is the state vector, $u \in \mathbb{R}^{n_u}$ is the input vector, and $y \in \mathbb{R}^{n_y}$ is the output. The disturbance vector $d \in \mathbb{R}^{n_d}$ is assumed to be known/measurable at training stage, but unknown in online operations. The terms $z = z(x, u, d)$, $f = f(x, u, d)$, and $r = r(x, u, d)$ are the *gates* of the network, which are defined as:

$$\begin{aligned} z &= \sigma(W_z u + V_z d + U_z x + b_z), \\ f &= \sigma(W_f u + V_f d + U_f x + b_f), \\ r &= \phi(W_r u + V_r d + U_r x + b_r). \end{aligned} \quad (1b)$$

The set of learnable parameters of the model (weights) is

$$\Phi = \{W_z, W_f, W_r, V_z, V_f, V_r, V_o, U_z, U_f, U_r, U_o, b_z, b_f, b_r, b_o\}.$$

Assumption 1: The input u of the GRU and the disturbance d are unity bounded, *i.e.*,

$$\begin{aligned} u \in \mathcal{U} &\triangleq \{u \in \mathbb{R}^{n_u} : \|u\|_\infty \leq 1\}, \\ d \in \mathcal{D} &\triangleq \{d \in \mathbb{R}^{n_d} : \|d\|_\infty \leq 1\}. \end{aligned}$$

Note that Assumption 1 is common when operating with RNN models, and can be satisfied by normalizing the input and disturbance data, provided that they are limited [14].

A. Stability properties

In this section, we analyze the stability properties of the GRU model (1). As the properties considered are usually regional, let us take the following mild assumption and lemma.

Assumption 2: The GRU model (1) is initialized within an arbitrarily large, but finite, set \mathcal{X} defined as

$$\mathcal{X} \triangleq \{x \in \mathbb{R}^{n_x} : \|x\|_\infty \leq \tilde{x}\}, \quad \text{with } \tilde{x} \geq 1.$$

Lemma 1 (Invariant set of GRUs [15]): The set \mathcal{X} is an invariant set for the GRU model (1). That is, $x \in \mathcal{X} \implies x^+ = f_\Sigma(x, u, d) \in \mathcal{X}$, for any $u \in \mathcal{U}$ and $d \in \mathcal{D}$.

Proof: The proof of [15, Lemma 3.3] can be applied to (1) by considering the augmented input $\tilde{u} = [u^\top, d^\top]^\top$. ■

In light of Assumption 1 and Lemma 1, the following bounds can be provided on the gates z , f , and r .

Lemma 2: The gates z , f , and r of the GRU model (1) can be bounded, for any component $j \in \{1, \dots, n_x\}$, as follows

$$\begin{aligned} 0 &< 1 - \check{\sigma}_z \leq z^{(j)} \leq \check{\sigma}_z < 1, \\ 0 &< 1 - \check{\sigma}_f \leq f^{(j)} \leq \check{\sigma}_f < 1, \\ -1 &< -\check{\phi}_r \leq r^{(j)} \leq \check{\phi}_r < 1, \end{aligned} \quad (2)$$

where the bounds are defined as

$$\begin{aligned} \check{\sigma}_z &= \sigma(\|\tilde{W}_z \ U_z \tilde{x} \ b_z\|_\infty), \\ \check{\sigma}_f &= \sigma(\|\tilde{W}_f \ U_f \tilde{x} \ b_f\|_\infty), \\ \check{\phi}_r &= \phi(\|\tilde{W}_r \ U_r \tilde{x} \ b_r\|_\infty), \end{aligned} \quad (3)$$

with $\tilde{W}_\nu = [W_\nu, V_\nu]$ for $\nu \in \{r, f, z\}$.

Proof: The proof of [15, Lemma 3.5] can be applied to (1) by considering the augmented input $\tilde{u} = [u^\top, d^\top]^\top$. ■

At this stage, the following stability property is introduced.

Definition 1 (δ ISS): The model (1) is said to be regionally Incrementally Input-to-State Stable (δ ISS) if there exists functions $\beta \in \mathcal{KL}$ and $\gamma_u, \gamma_d \in \mathcal{K}_\infty$ such that, for any pairs of initial conditions $x_0^a, x_0^b \in \mathcal{X}$, input sequences $u_{0:k}^a, u_{0:k}^b \in \mathcal{U}$, and disturbance sequences $d_{0:k}^a, d_{0:k}^b \in \mathcal{D}$, it holds:

$$\begin{aligned} \|x^a - x^b\|_2 &\leq \beta(\|x_0^a - x_0^b\|_2, k) \\ &+ \gamma_u \left(\max_{\tau \in \{0, \dots, k-1\}} \|u_\tau^a - u_\tau^b\|_2 \right) + \gamma_d \left(\max_{\tau \in \{0, \dots, k-1\}} \|d_\tau^a - d_\tau^b\|_2 \right), \end{aligned}$$

where x^ν , with $\nu \in \{a, b\}$, denotes the state of model (1) initialized in x_0^ν and fed with $u_{0:k-1}^\nu$ and $d_{0:k-1}^\nu$.

The stability property in Definition 1 entails that the state trajectories are asymptotically independent of the initial conditions. Therefore, the following theorem can be stated.

Theorem 1 (δ ISS of the GRU model): A sufficient condition for the δ ISS of the GRU model (1), in the sense specified by Definition 1 and with $\check{\sigma}_z, \check{\sigma}_f, \check{\phi}_r$ defined as in Lemma 2, is the following:

$$\|U_r\|_\infty \left(\frac{1}{4} \check{\tilde{x}} \|U_f\|_\infty + \check{\sigma}_f \right) + \frac{1}{4} \frac{(\check{\tilde{x}} + \check{\phi}_r)}{(1 - \check{\sigma}_z)} \|U_z\|_\infty - 1 < 0. \quad (4)$$

Proof: The proof is reported in Appendix A. ■

Along the lines of [6], the devised δ ISS condition (4) can be enforced during the model's training procedure, thus learning a provenly- δ ISS GRU model exploiting input-output data. It is worth noting that disturbance historical data are supposed to be available for the training procedure. In the remainder of this paper, we assume that the GRU model (1) has been trained as described, *e.g.*, in [6], and that (4) is satisfied by the trained weights Φ^* .

Assumption 3: The model (1) is δ ISS with respect to the sets \mathcal{X} , \mathcal{D} and \mathcal{U} , *i.e.*, condition (4) holds true.

III. CONTROL STRATEGY

For the system learned by the GRU model (1), we propose a control strategy called *State & Output Disturbance* (SOD) that involves to design a state observer that estimates both the state vector x and the disturbance vector d , which is assumed to be non measurable online. The resulting estimations (\hat{x}, \hat{d}) are later used to initialize the NMPC model, which yields the optimal input u^* that is applied to the plant (see Fig. 1).

A. Observer design

The state observer is designed under the common assumption that the disturbance d is piecewise-constant:

$$d^+ = d. \quad (5)$$

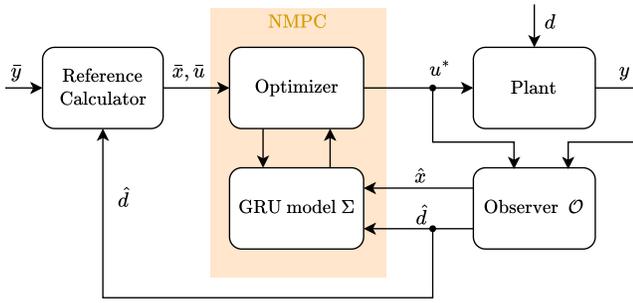


Fig. 1. Scheme of the proposed NMPC control approach.

For the enlarged model, achieved by augmenting the GRU model (1) with the disturbance dynamics (5), we propose the following Luenberger-like state observer:

$$\mathcal{O}(\Phi_o) : \begin{cases} \hat{x}^+ = f_o(\hat{x}, \hat{d}, u, y; \Phi_o) = \hat{z} \circ \hat{x} + (1 - \hat{z}) \circ \hat{r}, \\ \hat{d}^+ = h_o(\hat{x}, \hat{d}, y; \Phi_o) = \text{sat}_{\mathcal{D}}(\hat{d} + L_d(y - \hat{y})), \\ \hat{y} = g_o(\hat{x}, \hat{d}; \Phi_o) = V_o \hat{d} + U_o \hat{x} + b_o, \end{cases} \quad (6a)$$

where

$$\begin{aligned} \hat{z} &= \sigma(W_z u + V_z \hat{d} + U_z \hat{x} + b_z + L_z(y - \hat{y})), \\ \hat{f} &= \sigma(W_f u + V_f \hat{d} + U_f \hat{x} + b_f + L_f(y - \hat{y})), \\ \hat{r} &= \phi(W_r u + V_r \hat{d} + U_r \hat{f} \circ \hat{x} + b_r), \end{aligned} \quad (6b)$$

and $\text{sat}_{\mathcal{D}}(\cdot)$ denotes the element-wise saturation to the box \mathcal{D} as in [16], ensuring that the disturbance estimate remains consistent to $\hat{d} \in \mathcal{D}$ also during transients. The matrices $L_z, L_f \in \mathbb{R}^{n_x \times n_y}$ and $L_d \in \mathbb{R}^{n_d \times n_y}$ are the gains of the observer to be designed and, together with the already trained weights of the model (1), represent the observer parameters $\Phi_o = \Phi^* \cup \{L_z, L_f, L_d\}$. The observer requires a suitable tuning of its gains to compute its future state \hat{x}^+ and disturbance \hat{d}^+ estimates at any time instant. The design must also guarantee the convergence of both estimations with their true values, as specified in the following definition.

Definition 2: The observer $\mathcal{O}(\Phi_o)$ is said to be *convergent* if, for any initial state $x_0 \in \mathcal{X}$ and any disturbance $d_0 \in \mathcal{D}$, given the applied input sequence $u_{0:k}$ and measured output sequence $y_{0:k}(x_0, d_0, u_{0:k}; \Phi^*)$, the state estimate $\hat{x}(\hat{x}_0, \hat{d}_0, u_{0:k}, y_{0:k}; \Phi_o)$ and the disturbance estimate $\hat{d}(\hat{x}_0, \hat{d}_0, y_{0:k}; \Phi_o)$ converge, respectively, to their true values for any initial guess $\hat{x}_0 \in \mathcal{X}$ and $\hat{d}_0 \in \mathcal{D}$. That is, there exists a function $\beta_o(\cdot, k) \in \mathcal{KL}$ such that

$$\begin{bmatrix} \|x - \hat{x}\| \\ \|d - \hat{d}\|_2 \end{bmatrix} \leq \beta_o \left(\begin{bmatrix} \|x_0 - \hat{x}_0\| \\ \|d_0 - \hat{d}_0\|_2 \end{bmatrix}, k \right). \quad (7)$$

Theorem 2 (Observer's exponential convergence): A sufficient condition for the exponential convergence of the observer (6) is that L_z, L_f, L_d are such that the matrix

$$\bar{A} = \bar{A}(z, L_z, L_f, L_d) = \begin{bmatrix} \kappa_{11} & \kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{bmatrix}, \quad (8)$$

where, for all $z \in [1 - \check{\sigma}_z, \check{\sigma}_z]$,

$$\begin{aligned} \kappa_{11} &= z + (1 - z) \left(\frac{1}{4} \check{x} \|U_f - L_f U_o\|_{\infty} + \check{\sigma}_f \right) \|U_r\|_{\infty} \\ &\quad + \frac{1}{4} (\check{x} + \check{\phi}_r) \|U_z - L_z U_o\|_{\infty}, \\ \kappa_{12} &= (1 - z) \left(\frac{1}{4} \check{x} \|U_r\|_{\infty} \|V_f - L_f V_o\|_{\infty} + \|V_r\|_{\infty} \right) \\ &\quad + \frac{1}{4} (\check{x} + \check{\phi}_r) \|V_z - L_z V_o\|_{\infty}, \\ \kappa_{21} &= \|L_d U_o\|_{\infty}, \\ \kappa_{22} &= \|I - L_d V_o\|_{\infty}, \end{aligned}$$

is Schur stable, *i.e.*, its spectral radius is $\rho(\bar{A}) < 1$.

Proof: The proof is reported in Appendix B. \blacksquare

Thus the problem of designing the observer (6) amounts to finding gains L_z, L_f, L_d that satisfy Theorem 2. This can be done by formulating the observer design problem as a nonlinear optimization problem, as discussed below.

Proposition 1 (Optimal observer design): The gains of the GRU observer (6) that allow to address Theorem 2 can be found by solving:

$$\begin{aligned} L_z^*, L_f^*, L_d^* &= \arg \min_{\{L_z, L_f, L_d\}} \|\bar{A}\|_2 \quad (9a) \\ \text{s.t.} \quad -1 + \kappa_{11} + \kappa_{22} &< \kappa_{11} \kappa_{22} - \kappa_{12} \kappa_{21} < 1 \quad (9b) \end{aligned}$$

with the constraint evaluated in $z = \check{\sigma}_z$ and $z = 1 - \check{\sigma}_z$.

Proof: In light of [7, Lemma 1], \bar{A} is Schur stable if and only if $-1 + \kappa_{11} + \kappa_{22} < \kappa_{11} \kappa_{22} - \kappa_{12} \kappa_{21} < 1$ for all $z \in [1 - \check{\sigma}_z, \check{\sigma}_z]$. Since κ_{11} and κ_{12} are linear with respect to z , while κ_{21} and κ_{22} are independent of z , it is enough to impose constraint (9b) for the extreme values $z = \check{\sigma}_z$ and $z = 1 - \check{\sigma}_z$. By [7, Lemma 1], any feasible solution of (9) satisfies Theorem 2. \blacksquare

Note that since $\rho(\bar{A}) \leq \|\bar{A}\|_2$, the minimization of $\|\bar{A}\|_2$ promotes a faster worst-case convergence of the estimates. The fulfillment of Theorem 2 is guaranteed by constraint (9b), which is equivalent to $\rho(\bar{A}) < 1$, see [7, Lemma 1].

B. NMPC formulation

A nonlinear MPC law based on the GRU model (1) is proposed. We denote by $u_{0:N-1} = \{u_0, \dots, u_{N-1}\}$ the input sequence applied throughout the prediction horizon N , and by $x_{0:N} = \{x_0, \dots, x_N\}$ the corresponding predicted state trajectory, where x_n and u_n are respectively the state and input predicted at step $k + n$. Moreover, the controller considers a setpoint \bar{y} that meets the following assumption.

Assumption 4: For an output setpoint $\bar{y} \in \mathcal{Y} \subseteq \mathbb{R}^{n_y}$ and any disturbance $\hat{d} \in \mathcal{D}$, there exists a feasible equilibrium $\bar{u} \in \mathcal{U}$ and $\bar{x} \in \mathcal{X}$ such that $\bar{x} = f_{\Sigma}(\bar{x}, \bar{u}, \hat{d})$ and $\bar{y} = g_{\Sigma}(\bar{x}, \hat{d})$.

At each time step k , the control objective is to minimize:

$$J(\cdot, \cdot) = \sum_{n=0}^{N-1} \left(\|x_n - \bar{x}\|_Q^2 + \|u_n - \bar{u}\|_R^2 \right) + V(x_N, \bar{x}, \bar{u}),$$

where \bar{x} and \bar{u} are steady-state values of x and u such that, for the constant disturbance $d = \hat{d}$, the output is equal to the setpoint \bar{y} . In $J(\cdot, \cdot)$, the first term is a stage cost with weighting matrices $Q \succ 0$ and $R \succ 0$ penalizing, throughout the prediction horizon, the deviation of the predicted state and input trajectories from the references \bar{x} and \bar{u} . The last term

is a terminal cost commonly designed to ensure closed-loop stability of the MPC. Following [8], it is defined as

$$V(x_N, \bar{x}, \bar{u}) = \sum_{m=0}^M \|x_{N+m} - \bar{x}\|_S^2,$$

where x_{N+m} with $m \in \{0, \dots, M\}$ denotes the state prediction from step N to $N + M$ under the constant input $u = \bar{u}$. For the design of a terminal horizon M and a weight matrix S achieving closed-loop stability properties, see [8]. The optimal control problem is defined as:

$$u_{0:N-1}^* = \arg \min_{u_{0:N-1}} J(x_{0:N}, u_{0:N-1}) \quad (10a)$$

$$\text{s.t. } x_0 = \hat{x}, \quad d_0 = \hat{d}, \quad (10b)$$

$$d_{n+1} = d_0, \quad (10c)$$

$$x_{n+1} = f_\Sigma(x_n, u_n, d_n), \quad (10d)$$

$$u_n \in \mathcal{U}, \quad (10e)$$

for all $n \in \{0, \dots, N - 1\}$. By constraint (10b), the model is initialized by \hat{x} and \hat{d} , *i.e.*, the current state and disturbance estimates given by the observer (6). The disturbance is considered constant throughout the prediction horizon by (10c), and the dynamics of the model is imposed by constraint (10d). The control variable is constrained within the set \mathcal{U} via (10e), see Assumption 1.

At each time step k , the sequence $u_{0:N-1}^*$ is obtained by solving problem (10), but only the first input u_0^* is applied to the plant. Algorithm 1 summarizes the steps of our proposed NMPC controller based on the SOD model.

Algorithm 1

Input: Setpoint \bar{y} .

- 1: Estimate the state \hat{x} and disturbances \hat{d} via the GRU observer (6) given the output measurement y .
 - 2: Calculate the equilibrium (\bar{x}, \bar{u}) for the output setpoint \bar{y} and the estimated disturbance \hat{d} subject to the GRU dynamics such that $\bar{x} = f_\Sigma(\bar{x}, \bar{u}, \hat{d})$ and $\bar{y} = g_\Sigma(\bar{x}, \hat{d})$.
 - 3: Solve (10) to obtain the control sequence $u_{0:N}^*$.
 - 4: Apply the first control action u_0^* to the plant.
-

Remark 1: The stability of the adopted MPC algorithm is ensured in the presence of constant disturbances. If d is piecewise-constant, the disturbance estimate \hat{d} will converge, see [17], so that the stability is recovered.

Remark 2: To deal with disturbances in the MPC design, it is common to consider a model where disturbances act linearly only on the output and are estimated as previously described. In this approach, henceforth named *Output Disturbance* (OD) estimation, the disturbance is intended to include not only the effect of exogenous signals, but also unmodeled dynamics. Although the OD approach usually provides suitable control results, it does not account for previous system information, *e.g.*, the presence of disturbances acting on the state equation, as done in physics-based approaches. Moreover, the estimated disturbance is dissimilar from the real one.

IV. CASE STUDY

The control strategy is evaluated in the pH neutralization process [18] illustrated in Fig. 2. This system is composed

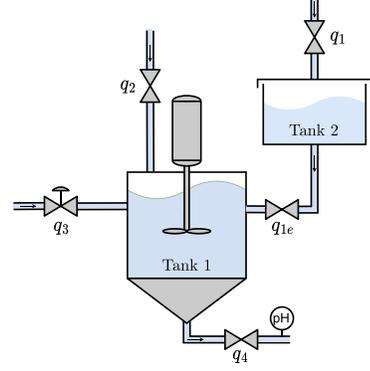


Fig. 2. Scheme of the pH neutralization system.

of two tanks: the reactor tank #1, which is fed with the flow rates q_{1e} (acid flow), q_2 (buffer flow) and q_3 (alkaline base flow), and the tank #2, which is fed with the acid flow rate q_1 and provides as output q_{1e} . It is assumed that $q_{1e} = q_1$ as the hydraulic dynamics of the tank #2 is much faster than the others involved. The flow rates q_1 and q_2 cannot be manipulated, whereas q_3 can be regulated by a valve and represents the control action u . In particular, we assume that $d = q_1$ is a piecewise-constant disturbance while q_2 is fixed to the nominal value. The pH is measured from the constant flow rate q_4 at the output of the reactor tank. The system model is represented by nonlinear third-order dynamics given by the equations and parameters detailed in [18]. The objective is to regulate the pH value to a given setpoint $\bar{y} = 8$ considering input constraints $[u^{\min}, u^{\max}] = [12.5, 17]$ mL/s.

To analyze the advantages of the proposed SOD controller (NMPC_{SOD}), we compare it with the OD approach (NMPC_{OD}) outlined in Remark 2. In both approaches, the system has been identified with a single-layer GRU model with $n_x = 7$ units that enjoys the exponential δ ISS property, following the procedure detailed in [6]. The datasets to learn the models are obtained from a plant simulator operated for approx. 42 h with a sampling time $T_s = 15$ s. In particular,

- The GRU_{SOD} model (1) is trained using a dataset of input, output and disturbance trajectories: $\mathcal{T}_{\text{SOD}} = \{u_{0:T}, d_{0:T}, y_{0:T}\}$,
- The GRU_{OD} model, which is defined as $x^+ = f_\Sigma(x, u; \Phi) + d$, and $y = g_\Sigma(x; \Phi) + d$, is trained with $d = 0$, thus the dataset employed is $\mathcal{T}_{\text{OD}} = \{u_{0:T}, y_{0:T}\}$.

Note that these trajectories must be normalized to be used for the training procedures but, in the figures, they are depicted in their original scales for the sake of interpretability. The corresponding datasets are partitioned into a training set ($T_{\text{tr}} = 70\% T$), a validation set ($T_{\text{val}} = 15\% T$), and an independent test set ($T_{\text{te}} = 15\% T$) to assess the performance of the learned models. The training procedures are carried out minimizing, with the RMSProp optimizer [14], the truncated simulation error of the model, plus a regularization term to enforce the model δ ISS [6].

At the end of the training, the models' accuracy are quantified on the independent test sets via the FIT index. This index is the ratio between the open-loop simulation error on the test set, after a washout period T_w , and the root

TABLE I
FIT INDEX IN THE TEST SETS FOR BOTH MODELS

	FIT in $T_{te,OD}$	FIT in $T_{te,SOD}$
GRU _{OD}	96.56 %	77.28 %
GRU _{SOD}	96.04 %	95.96 %

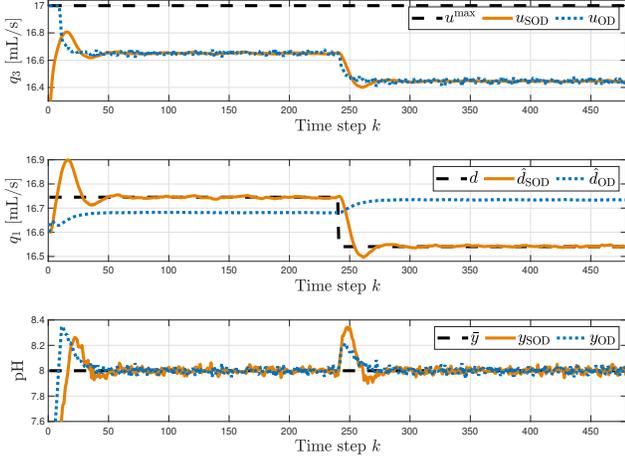


Fig. 3. Results on implementing NMPC_{SOD} and NMPC_{OD} approaches.

mean square value of the output w.r.t its mean $y_{te,avg}$:

$$FIT(\%) = 100 \cdot \left(1 - \frac{\sum_{k=T_w}^{T_{te}} \|y(x_0, d_{te,0:k}, u_{te,0:k}; \Phi^*) - y_{te,k}\|_2}{\sum_{k=T_w}^{T_{te}} \|y_{te,k} - y_{te,avg}\|_2} \right).$$

The learned δ ISS GRU_{SOD} model scores a 95.96% FIT value in the test set $T_{te,SOD}$ compared to the ground truth. Comparing the FIT index of both models for two test sets with disturbances $T_{te,SOD}$ and without disturbances $T_{te,OD}$ (see Table I), the GRU_{SOD} model is more accurate because it leverages disturbances knowledge during the training stage.

The observer's gains (6) are computed by solving the problem (9) but, due to the conservativity of the approach, a fine tuning has been implemented to improve performance. Note that the GRU_{OD} model is augmented with a disturbance model as (5) and the observer $\mathcal{O}_{OD}(\Phi_o) \triangleq \{\hat{x}^+ = f_o(\hat{x}, u, y; \Phi_o), \hat{d}^+ = h_o(\hat{x}, \hat{d}, y; \Phi_o), \hat{y} = g_o(\hat{x}, \hat{d}; \Phi_o)\}$ is designed with the gains resulting from solving a problem analogous to (9) to estimate its states and disturbances.

The simulation length is 2 h. To design the MPC parameters of both approaches, we consider the prediction horizon $N = 20$, the terminal horizon $M = 20$, and the weighting matrices $Q = I_{n_x \times n_x}$, $S = 2Q$, and $R = 0.25$.

Fig. 3 depicts the input, the disturbance, and the output trajectories by implementing both control approaches. Note that the output measurements y are corrupted by white noise ($\sigma = 0.02$). As shown, both controllers lead the output to the setpoint while respecting the input constraints. Nonetheless, there are some differences we would like to remark on. With NMPC_{OD}, the noise from the output measurement is reflected in the input signal, e.g., the standard deviation of $u_{OD,60:200}$ is $11.37 \cdot 10^{-3}$. Conversely, the control action in NMPC_{SOD} is filtered from that noise, as shown in the top graph of Fig. 3 (the standard deviation of $u_{SOD,60:200}$ is $3.28 \cdot 10^{-3}$), which is beneficial to avoid the wear of the valve controlling

q_3 . Additionally, NMPC_{SOD} provides the convergence of the disturbance estimate (second plot in Fig. 3), obtaining a suitable online estimate of this unmeasurable input that affects the system. In contrast, the disturbance estimated with NMPC_{OD} does not represent a real disturbance as $d = q_1$, but rather a measure of uncertainties to regulate the output.

V. CONCLUSIONS

This work investigates online estimation of disturbances in states and outputs of Gated Recurrent Unit (GRU) neural networks, which are trained to describe the system dynamics with input-to-state stability. A Luenberger-like observer with convergence guarantees is proposed for the estimation of disturbances and the black-box states of the GRU model. Finally, we design an offset-free nonlinear MPC for setpoint tracking and we test it in a nonlinear benchmark. The results show some improvements over the usual approach in which disturbances are considered to act linearly only on the output. Remarkably, our approach provides a reliable disturbance estimate, crucial for optimizing control and developing robust monitoring and fault detection protocols.

Future work will aim to develop MPC methods that avoid recalculating state and input references for each new disturbance estimate. Additionally, we will also explore robustness guarantees to handle transient variations between estimated and actual disturbances.

APPENDIX

A. Proof of Theorem 1

Let x^ν be the state of the GRU model (1), initialized in $x_0^\nu \in \mathcal{X}$ and fed with the input sequence $u_{0:k}^\nu \in \mathcal{U}$ and by the disturbance $d_{0:k}^\nu \in \mathcal{D}$, for both trajectories $\nu \in \{a, b\}$. Consider the augmented input vector $\tilde{u}^\nu = [u^{\nu\top}, d^{\nu\top}]^\top \in \tilde{\mathcal{U}}$, where $\tilde{\mathcal{U}} = \mathcal{U} \times \mathcal{D}$ and let $\Delta x = x^a - x^b$, $\Delta x_0 = x_0^a - x_0^b$, and $\Delta \tilde{u} = \tilde{u}^a - \tilde{u}^b$. Following a procedure similar to [15, Proof of Theorem 3.8] one can bound Δx as follows

$$\begin{aligned} \|\Delta x\|_\infty &\leq \lambda_\delta^k \|\Delta x_0\|_\infty + \frac{\tilde{\kappa} \Delta \tilde{u}}{1 - \lambda_\delta} \max_{\tau \in \{1, \dots, k-1\}} \|\Delta u_\tau\|_\infty \\ &\quad + \frac{\tilde{\kappa} \Delta \tilde{u}}{1 - \lambda_\delta} \max_{\tau \in \{1, \dots, k-1\}} \|\Delta d_\tau\|_\infty. \end{aligned}$$

By [15, Lemma 2.1], the system is δ ISS (in the sense specified by Definition 1) with functions $\beta(s, k) = \sqrt{n_x} \lambda_\delta^k s$ and $\gamma_u(s) = \gamma_d(s) = \frac{\sqrt{n_x} \tilde{\kappa} \Delta \tilde{u}}{1 - \lambda_\delta} s$. ■

B. Proof of Theorem 2

Consider the GRU initialized at $x_0 \in \mathcal{X}$, fed with (constant) disturbance $d_0 \in \mathcal{D}$ and input $u_{0:k}$. Let denote the resulting (unknown) state at k as $x = x(x_0, d_0, u_{0:k}; \Phi^*)$ and the measured output as $y = y(x_0, d_0, u_{0:k}; \Phi^*)$. Let $\hat{x} = \hat{x}(\hat{x}_0, \hat{d}_0, u_{0:k}, y_{0:k}; \Phi_o)$ and $\hat{d} = \hat{d}(\hat{x}_0, \hat{d}_0, u_{0:k}, y_{0:k}; \Phi_o)$ be the state and disturbance estimates, respectively, yielded by observer (6) with initial guess $\hat{x}_0 \in \mathcal{X}$ and $\hat{d}_0 \in \mathcal{D}$, and fed by $u_{0:k} \in \mathcal{U}_{0:k}$ and $y_{0:k}$. First, consider the j -th component of the state error $x^+ - \hat{x}^+$. From (6) and (1), it follows

$$\begin{aligned} x_{(j)}^+ - \hat{x}_{(j)}^+ &= z_{(j)} x_{(j)} + (1 - z_{(j)}) r_{(j)} - \hat{z}_{(j)} \hat{x}_{(j)} - (1 - \hat{z}_{(j)}) \hat{r}_{(j)} \\ &= z_{(j)} (x_{(j)} - \hat{x}_{(j)}) + (z_{(j)} - \hat{z}_{(j)}) \hat{x}_{(j)} \\ &\quad + (1 - z_{(j)}) (r_{(j)} - \hat{r}_{(j)}) + (z_{(j)} - \hat{z}_{(j)}) \hat{r}_{(j)}. \end{aligned}$$

Then, take the absolute value and recall $z_{(j)}, \hat{z}_{(j)} \in (0, 1)$

$$\begin{aligned} |x_{(j)}^+ - \hat{x}_{(j)}^+| &\leq z_{(j)}|x_{(j)} - \hat{x}_{(j)}| + |z_{(j)} - \hat{z}_{(j)}|\hat{x}_{(j)} \\ &\quad + (1 - z_{(j)})|r_{(j)} - \hat{r}_{(j)}| + |z_{(j)} - \hat{z}_{(j)}|\hat{r}_{(j)}. \end{aligned} \quad (11)$$

Since $\hat{d} \in \mathcal{D}$, \mathcal{X} is also an invariant set for \hat{x} , implying that

$$|\hat{x}_{(j)}| \leq \|\hat{x}\|_\infty \leq \check{x}, \quad (12)$$

and, as in Lemma 2, the term $\hat{r}_{(j)}$ can be bounded over \mathcal{X} as

$$|\hat{r}_{(j)}| \leq \|\hat{r}\|_\infty \leq \check{\phi}_r.$$

Recalling the linearity of the output transformation and that $\sigma(\cdot)$ is a $\frac{1}{4}$ -Lipschitz function, we have

$$\begin{aligned} |z_{(j)} - \hat{z}_{(j)}| &\leq \|z - \hat{z}\|_\infty \\ &\leq \frac{1}{4}\|U_z - L_z U_o\|_\infty \|x - \hat{x}\|_\infty + \frac{1}{4}\|V_z - L_z V_o\|_\infty \|d - \hat{d}\|_\infty \end{aligned}$$

and, similarly,

$$\begin{aligned} |f_{(j)} - \hat{f}_{(j)}| &\leq \|f - \hat{f}\|_\infty \\ &\leq \frac{1}{4}\|U_f - L_f U_o\|_\infty \|x - \hat{x}\|_\infty + \frac{1}{4}\|V_f - L_f V_o\|_\infty \|d - \hat{d}\|_\infty. \end{aligned}$$

Since $\phi(\cdot)$ is a 1-Lipschitz, $|r_{(j)} - \hat{r}_{(j)}|$ can be bounded as

$$\begin{aligned} |r_{(j)} - \hat{r}_{(j)}| &\leq \|r - \hat{r}\|_\infty \\ &\leq 1\|U_r[(f - \hat{f})\hat{x} + f(x - \hat{x})] + V_r(d - \hat{d})\|_\infty \\ &\stackrel{(2),(12)}{\leq} \|U_r\|_\infty \left[\check{x}\|f - \hat{f}\|_\infty + \check{\sigma}_f\|x - \hat{x}\|_\infty \right] \\ &\quad + \|V_r\|_\infty \|d - \hat{d}\|_\infty, \end{aligned}$$

then considering the bound of $|f_{(j)} - \hat{f}_{(j)}|$ and reordering:

$$\begin{aligned} |r_{(j)} - \hat{r}_{(j)}| &\leq \|U_r\|_\infty \left(\frac{1}{4}\check{x}\|U_f - L_f U_o\|_\infty + \check{\sigma}_f \right) \|x - \hat{x}\|_\infty \\ &\quad + \left(\frac{1}{4}\check{x}\|U_r\|_\infty \|V_f - L_f V_o\|_\infty + \|V_r\|_\infty \right) \|d - \hat{d}\|_\infty. \end{aligned}$$

In light of previous bounds, the inequality (11) is bounded as

$$\begin{aligned} |x_{(j)}^+ - \hat{x}_{(j)}^+| &\leq \kappa_{11}\|x - \hat{x}\|_\infty + \kappa_{12}\|d - \hat{d}\|_\infty, \text{ where} \\ \kappa_{11} &= z_{(j)} + (1 - z_{(j)})\|U_r\|_\infty \left(\frac{1}{4}\check{x}\|U_f - L_f U_o\|_\infty + \check{\sigma}_f \right) \\ &\quad + \frac{1}{4}(\check{x} + \check{\phi}_r)\|U_z - L_z U_o\|_\infty, \\ \kappa_{12} &= (1 - z_{(j)}) \left(\frac{1}{4}\check{x}\|U_r\|_\infty \|V_f - L_f V_o\|_\infty + \|V_r\|_\infty \right) \\ &\quad + \frac{1}{4}(\check{x} + \check{\phi}_r)\|V_z - L_z V_o\|_\infty. \end{aligned}$$

Concerning the disturbance observation error, consider the j -th component of the difference between (1) and (6):

$$\begin{aligned} d_{(j)}^+ - \hat{d}_{(j)}^+ &= d_{(j)} - \text{sat}_{\mathcal{D}}(\hat{d}_{(j)} + L_{d_j}(y_{(j)} - \hat{y}_{(j)})) \\ &\stackrel{d \in \mathcal{D}}{=} \text{sat}_{\mathcal{D}}(d_{(j)}) - \text{sat}_{\mathcal{D}}(\hat{d}_{(j)} + L_{d_j}(y_{(j)} - \hat{y}_{(j)})), \end{aligned}$$

where L_{d_j} denotes the j -th row of L_d . Then we take the absolute value, recalling that the $\text{sat}_{\mathcal{D}}(\cdot)$ is 1-Lipschitz and that, by definition, $|d_{(j)}| \leq \|d\|_\infty$, which yields

$$\begin{aligned} |d_{(j)}^+ - \hat{d}_{(j)}^+| &\leq \left| \text{sat}_{\mathcal{D}}(d_{(j)}) - \text{sat}_{\mathcal{D}}(\hat{d}_{(j)} + L_{d_j}(y_{(j)} - \hat{y}_{(j)})) \right| \\ &\leq \left\| d - \hat{d} - L_d(y - \hat{y}) \right\|_\infty \\ &\leq \underbrace{\|L_d U_o\|_\infty}_{\kappa_{21}} \|x - \hat{x}\|_\infty + \underbrace{\|I - L_d V_o\|_\infty}_{\kappa_{22}} \|d - \hat{d}\|_\infty. \end{aligned}$$

Finally, combining state and disturbances bounds, it follows

$$\begin{bmatrix} |x_{(j)}^+ - \hat{x}_{(j)}^+| \\ |d_{(j)}^+ - \hat{d}_{(j)}^+| \end{bmatrix} \leq \bar{A} \begin{bmatrix} \|x - \hat{x}\|_\infty \\ \|d - \hat{d}\|_\infty \end{bmatrix}, \quad (13)$$

where \bar{A} is defined as in (8). The Schur stability of \bar{A} thus implies the existence of $\mu_0 > 0$ and $\lambda_o \in (0, 1)$ such that

$$\left\| \begin{bmatrix} \|x - \hat{x}\|_\infty \\ \|d - \hat{d}\|_\infty \end{bmatrix} \right\|_2 \leq \mu_o \lambda_o^k \left\| \begin{bmatrix} \|x_0 - \hat{x}_0\|_\infty \\ \|d_0 - \hat{d}_0\|_\infty \end{bmatrix} \right\|_2, \quad (14)$$

therefore, (14) implies

$$\left\| \begin{bmatrix} x - \hat{x} \\ d - \hat{d} \end{bmatrix} \right\|_2 \leq \sqrt{n_x + n_d} \mu_o \lambda_o^k \left\| \begin{bmatrix} x_0 - \hat{x}_0 \\ d_0 - \hat{d}_0 \end{bmatrix} \right\|_2. \quad (15)$$

The observer is therefore *exponentially convergent* with function $\beta_o(s, k) = \sqrt{n_x + n_d} \mu_o \lambda_o^k s$. ■

REFERENCES

- [1] L. Hewing, K. P. Wabersich, M. Menner, and M. N. Zeilinger, "Learning-based model predictive control: Toward safe learning in control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, pp. 269–296, 2020.
- [2] K. Zarzycki and M. Ławryńczuk, "LSTM and GRU neural networks as models of dynamical processes used in predictive control: A comparison of models developed for two chemical reactors," *Sensors*, vol. 21, no. 16, p. 5625, 2021.
- [3] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Proceedings of the International Conference on Machine Learning*, pp. 1310–1318, PMLR, 2013.
- [4] N. Lanzetti, Y. Z. Lian, A. Cortinovis, L. Dominguez, M. Mercangöz, and C. Jones, "Recurrent neural network based MPC for process industries," in *Proceedings of the 18th European Control Conference (ECC)*, pp. 1005–1010, IEEE, 2019.
- [5] E. Terzi, T. Bonetti, D. Saccani, M. Farina, L. Fagiano, and R. Scattolini, "Learning-based predictive control of the cooling system of a large business centre," *Control Engineering Practice*, vol. 97, 2020.
- [6] F. Bonassi, M. Farina, J. Xie, and R. Scattolini, "On recurrent neural networks for learning-based control: Recent results and ideas for future developments," *Journal of Process Control*, vol. 114, pp. 92–104, 2022.
- [7] E. Terzi, F. Bonassi, M. Farina, and R. Scattolini, "Learning model predictive control with long short-term memory networks," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 18, 2021.
- [8] F. Bonassi, A. La Bella, M. Farina, and R. Scattolini, "Nonlinear MPC design for incrementally ISS systems with application to GRU networks," *Automatica*, vol. 159, p. 111381, 2024.
- [9] G. Pannocchia, "Offset-free tracking MPC: A tutorial review and comparison of different formulations," in *Proceedings of the European Control Conference (ECC)*, pp. 527–532, IEEE, 2015.
- [10] L. Magni, G. De Nicolao, and R. Scattolini, "Output feedback and tracking of nonlinear systems with model predictive control," *Automatica*, vol. 37, no. 10, pp. 1601–1607, 2001.
- [11] M. Morari and U. Maeder, "Nonlinear offset-free model predictive control," *Automatica*, vol. 48, no. 9, pp. 2059–2067, 2012.
- [12] P. Tatjewski and M. Ławryńczuk, "Algorithms with state estimation in linear and nonlinear model predictive control," *Computers & Chemical Engineering*, vol. 143, p. 107065, 2020.
- [13] I. Schimperna and L. Magni, "Robust constrained nonlinear model predictive control with gated recurrent unit model," *Automatica*, vol. 161, p. 111472, 2024.
- [14] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.
- [15] F. Bonassi, *Reconciling deep learning and control theory: recurrent neural networks for model-based control design*. PhD thesis, 2023.
- [16] I. Schimperna and L. Magni, "Robust offset-free constrained model predictive control with long short-term memory networks—Extended version," *ArXiv preprint arXiv:2303.17304*, 2023.
- [17] P. O. Scokaert, J. B. Rawlings, and E. S. Meadows, "Discrete-time stability with perturbations: Application to model predictive control," *Automatica*, vol. 33, no. 3, pp. 463–470, 1997.
- [18] M. A. Henson and D. E. Seborg, "Adaptive nonlinear control of a pH neutralization process," *IEEE Transactions on Control Systems Technology*, vol. 2, no. 3, pp. 169–182, 1994.