*IMA Journal of Mathematical Control and Information* Page 1 of 16
doi:10.1093/imamci/dnnxxx

# Local output feedback stabilization of reaction-diffusion PDEs with saturated measurement*

HUGO LHACHEMI[1] AND CHRISTOPHE PRIEUR[2]
[1] *Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systèmes, 91190, Gif-sur-Yvette, France.*
*(e-mail: hugo.lhachemi@centralesupelec.fr).*
[2] *Université Grenoble Alpes, CNRS, Grenoble-INP, GIPSA-lab, F-38000, Grenoble, France.*
*(e-mail: christophe.prieur@gipsa-lab.fr)*

This paper addresses the topic of output feedback stabilization of general 1-D reaction-diffusion PDEs in the presence of a saturation in the measurement. The boundary control and the second boundary condition take the form of Dirichlet/Neumann/Robin boundary conditions. The measurement is selected as a boundary Dirichlet trace. The boundary measurement, as available for feedback control, is assumed to be subject to a saturation. In this context, we achieve the local exponential stabilization of the reaction-diffusion PDE while estimating a subset of the domain of attraction of the origin.

*Keywords*: reaction-diffusion PDE, output feedback, saturated measurement, local stabilization.

## 1. Introduction

Due to their physical limitations, actuators and sensors are inherently subject to saturation mechanisms. These saturation mechanisms introduce stringent constraints on the design of control laws (Bernstein and Michel 1995, Hippe 2006). Even in the most favorable case of finite-dimensional linear time invariant (LTI) systems, saturation mechanisms introduce harmful nonlinear phenomena characterized by multiple equilibrium points and bounded domains of attraction (Campo and Morari 1990). Since saturation mechanisms are ubiquitous in practical applications, the topic of feedback stabilization of finite-dimensional LTI systems despite the presence of input saturations has been intensively studied (Benzaouia et al. 2004, El Haoussi and Tissir 2007, Tarbouriech et al. 2011, Wei et al. 2014, Zaccarian and Teel 2011). In this context, one of the most fruitful approaches takes advantage of Lyapunov's direct method augmented with the use of a generalized sector condition (Tarbouriech et al. 2011, Lem. 1.6). This allows the derivation of 1) sufficient linear matrix inequality (LMI) conditions ensuring the local stability of the closed-loop plant; 2) a subset of the domain of attraction of the studied equilibrium point. A distinguished featured of the generalized sector condition is that this condition does not constrain the system trajectories so that the saturation mechanism is never active during the transient. Instead, the generalized sector condition can be used to compute initial conditions within the domain of attraction despite the actual activation of the saturation mechanism during the transient.

This paper is concerned with the feedback stabilization of infinite-dimensional systems, and particularly of partial differential equations (PDEs) (Liu et. al 2020, Zine and El Alami 2020), in the presence of a saturation mechanism. So far, this field of research has been mostly concerned with the case of input saturations. Such a problem was originally studied in (Slemrod 1989, Lasiecka and Seidman 2003) in the case of saturation mechanisms defined for control input functions evaluated in

*The work of H. Lhachemi has been partially supported by ANR PIA funding: ANR-20-IDEES-0002. The work of C. Prieur has been partially supported by MIAI@Grenoble Alpes (ANR-19-P3IA-0003)

the norm of an abstract functional space (typically the space of square integrable functions $L^2(0,1)$). Beside these seminal works, the topic of feedback stabilization of PDEs in the presence of input saturation mechanisms has been mainly focused on pointwise saturation mechanisms which are, in general, the most relevant for practical applications. The feedback stabilization of wave and Korteweg-de Vries PDEs under cone-bounded feedback and using Lyapunov's direct method have been extensively studied in (Prieur et al. 2016, Marx et al. 2017, Marx et al. 2017). The stabilization of reaction-diffusion PDEs in the presence of control input constraints was reported in (Dubljevic et al. 2006) by taking advantage of a model predictive control approach while singular perturbation techniques were studied in (Dubljevic et al. 2003). More recently, the feedback stabilization of reaction-diffusion PDEs using spectral reduction methods (Russell 1978, Coron and Trélat 2004, Coron and Trélat 2006, Lhachemi et al. 2019, Lhachemi and Prieur 2021) in the presence of an input saturation has been reported in (Mironchenko et al. 2021) in the case of a state-feedback with explicit estimation of the domain of attraction using LMIs. The extension to the local output feedback stabilization of a reaction-diffusion PDE using a either distributed or Dirichlet/Neumann boundary measurement, with also explicit estimation of the domain of attraction, was studied in (Lhachemi and Prieur 2022). It is worth noting that the state-feedback setting allows to achieve the local exponential stabilization of the the first modes of the system while preserving the stability of the residual infinite-dimensional dynamics (Mironchenko et al. 2021, Proposition 1). Hence, the region of attraction constrains only a finite number of modes of the initial condition (Mironchenko et al. 2021). In sharp contrast, the output feedback setting does not provide such a strong separation between the to-be-stabilized modes and the residual ones. This is because the measurement that is fed back to the controller is obtained based on the contribution of all the modes of the system. As a consequence, the region of attraction imposes constraints on all the modes of the initial condition (Lhachemi and Prieur 2022).

In this paper, we study for the first time the output feedback stabilization of reaction-diffusion PDEs in the presence of a saturated Dirichlet boundary measurement. Similarly to the case of an input saturation mechanism studied in (Lhachemi and Prieur 2022), the control strategy consists of a finite-dimensional controller (Curtain 1982, Sakawa 1983, Balas 1988, Harkort and Deutscher 2011, Sano 2012, Grüne and Meurer 2021) while the stability assessment is performed by combining Lyapunov's direct method and the above-mentioned generalized sector condition (Tarbouriech et al. 2011, Lem. 1.6). The controller architecture is composed of a finite-dimensional observer that leverages a control architecture reported first in (Sakawa 1983) along with a LMI-based approach introduced in (Katz and Fridman 2020). More precisely, we adopt the enhanced and general procedures described in (Lhachemi and Prieur 2022, Lhachemi and Prieur 2021) which allow the design, in a generic and systematic manner, of finite-dimensional observer-based control strategies for general 1-D reaction-diffusion PDEs, with Dirichlet/Neumann/Robin boundary control and Dirichlet/Neumann boundary measurement in a great variety of settings. These include regulation control (Lhachemi and Prieur 2021) and nonlinear boundary control (Lhachemi and Prieur 2021). This approach was shown to be efficient in the case of an input saturation in (Lhachemi and Prieur 2022). In this paper, we investigate the dual problem, namely the saturation of the measurement. It is worth noting that in the input saturation setting, the saturation applies to a signal that is only composed of a finite number of modes of the controller. In sharp contrast, the saturation of the measurement studied in this paper is more challenging because the saturation applies to a signal accounting for all the (infinite number and unmeasured) modes of the PDE system.

The paper is organized as follows. The problem description is reported in Section 2. The adopted control strategy is described in Section 3. The stability of the resulting closed-loop system is assessed in Section 4. The theoretical results are illustrated with some numerical simulations in Section 5. Concluding remarks are formulated in Section 6.

## 2. Notation, problem description, and spectral reduction

### 2.1 *Notation*

Real finite-dimensional spaces $\mathbb{R}^n$ are endowed with the Euclidean norm denoted by $\|\cdot\|$. The corresponding induced norms of matrices are also denoted by $\|\cdot\|$. For any vectors $x, y \in \mathbb{R}^n$, we note $x \leqslant y$ when each component of $x$ is less than or equal to the corresponding component of $y$. For any $x \in \mathbb{R}^n$, we note $|x|$ the vector of $\mathbb{R}^n$ obtained by substituting each component of $x$ by its absolute value. For any two vectors $X$ and $Y$ of arbitrary dimensions, we define $\mathrm{col}(X,Y) = [X^\top, Y^\top]^\top$. The space of square integrable functions on $(0,1)$ is denoted by $L^2(0,1)$ and is equipped with the inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)\,\mathrm{d}x$ while the associated norm is denoted by $\|\cdot\|_{L^2}$. For any given integer $m \geqslant 1$, the Sobolev space of order $m$ is denoted by $H^m(0,1)$ and is endowed with its usual norm $\|\cdot\|_{H^m}$. For a symmetric matrix $P \in \mathbb{R}^{n\times n}$, $P \succeq 0$ (resp. $P \succ 0$) means that $P$ is positive semi-definite (resp. positive definite) while $\lambda_M(P)$ (resp. $\lambda_m(P)$) denotes its maximal (resp. minimal) eigenvalue.

Let $(\phi_n)_{n\geqslant 1}$ be an arbitrarily given Hilbert basis of $L^2(0,1)$. For any two integers $1 \leqslant N < M$, we define the operators of projection:

$$\begin{aligned} \pi_N : \quad L^2(0,1) &\longrightarrow \mathbb{R}^N \\ f &\longmapsto \begin{bmatrix} \langle f,\phi_1\rangle & \ldots & \langle f,\phi_N\rangle \end{bmatrix}^\top \end{aligned}$$

and

$$\begin{aligned} \pi_{N,M} : \quad L^2(0,1) &\longrightarrow \mathbb{R}^{M-N} \\ f &\longmapsto \begin{bmatrix} \langle f,\phi_{N+1}\rangle & \ldots & \langle f,\phi_M\rangle \end{bmatrix}^\top . \end{aligned}$$

We finally define

$$\begin{aligned} \mathscr{R}_N : \quad L^2(0,1) &\longrightarrow L^2(0,1) \\ f &\longmapsto f - \sum_{n=1}^{N} \langle f,\phi_n\rangle\,\phi_n = \sum_{n\geqslant N+1} \langle f,\phi_n\rangle\,\phi_n . \end{aligned}$$

### 2.2 *Problem description*

We consider the reaction-diffusion equation described by

$$z_t(t,x) = (p(x)z_x(t,x))_x - \tilde{q}(x)z(t,x) \tag{2.1a}$$
$$\cos(\theta_1)z(t,0) - \sin(\theta_1)z_x(t,0) = 0 \tag{2.1b}$$
$$\cos(\theta_2)z(t,1) + \sin(\theta_2)z_x(t,1) = u(t) \tag{2.1c}$$
$$z(0,x) = z_0(x). \tag{2.1d}$$

for $t > 0$ and $x \in (0,1)$. Here we have $\theta_1, \theta_2 \in [0,\pi/2]$, $p \in \mathscr{C}^1([0,1])$ with $p > 0$, and $\tilde{q} \in \mathscr{C}^0([0,1])$. The boundary control input is $u(t) \in \mathbb{R}$, $z(t,\cdot) \in L^2(0,1)$ is the state of the reaction-diffusion PDE, and $z_0 \in H^2(0,1)$ with $\cos(\theta_1)z_0(0) - \sin(\theta_1)z_0'(0) = 0$ and $\cos(\theta_2)z_0(1) + \sin(\theta_2)z_0'(1) = u(0)$ is the initial condition. The system output available for feedback control is assumed to be a saturated left Dirichlet trace. More precisely, with $\theta_1 \in (0,\pi/2]$, the left Dirichlet trace is defined by:

$$y_D(t) = z(t,0). \tag{2.2}$$

Then the output that is available for feedback control is defined as

$$y_{D,\mathrm{sat}_l}(t) = \mathrm{sat}_l(y_D(t)) = \mathrm{sat}_l(z(t,0)) \tag{2.3}$$

for an arbitrary given level of saturation $l > 0$, where the saturation function $\mathrm{sat}_l : \mathbb{R} \to \mathbb{R}$ is defined as

$$\mathrm{sat}_l(y) = \begin{cases} y & \text{if } |y| \leqslant l; \\ l\dfrac{y}{|y|} & \text{if } |y| \geqslant l. \end{cases}$$

The objective is to design a finite-dimensional output feedback controller that achieves the local stabilization of (2.1) with a saturated measurement given by (2.3). Moreover, we aim at estimating a subset of the domain of attraction. To do so, it is classical to introduce the deadzone nonlinearity $\phi_l : \mathbb{R} \to \mathbb{R}$ defined for any $y \in \mathbb{R}$ by

$$\phi_l(y) = \mathrm{sat}_l(y) - y. \tag{2.4}$$

This representation is mainly motivated by the fact that this deadzone nonlinearity satisfies the following generalized sector condition borrowed from (Tarbouriech et al. 2011, Lem. 1.6) in the scalar case.

LEMMA 2.1 Let $l > 0$ be given. For any $y, \omega \in \mathbb{R}^m$ such that $|y - \omega| \leqslant l$ we have $\phi_l(y)(\phi_l(y) + \omega) \leqslant 0$.

### 2.3 *Preliminary spectral reduction*

2.3.1 *Properties of Sturm-Liouville operators* Reaction-diffusion PDEs are strongly related to Sturm-Liouville operators. We review here the definition of these operators along with their key properties that will be used in the sequel.

Let $\theta_1, \theta_2 \in [0, \pi/2]$, $p \in \mathscr{C}^1([0,1])$, and $q \in \mathscr{C}^0([0,1])$ with $p > 0$ and $q \geqslant 0$. The Sturm-Liouville operator is given by

$$\begin{array}{llll} \mathscr{A} : & D(\mathscr{A}) & \longrightarrow & L^2(0,1) \\ & f & \longmapsto & -(pf')' + qf \end{array} \tag{2.5}$$

with the domain of the operator defined by

$$\begin{aligned} D(\mathscr{A}) = \{ f \in H^2(0,1) : & \cos(\theta_1) f(0) - \sin(\theta_1) f'(0) = 0 \\ & \cos(\theta_2) f(1) + \sin(\theta_2) f'(1) = 0 \}. \end{aligned}$$

Then it holds that the eigenvalues $\lambda_n$, $n \geqslant 1$, of $\mathscr{A}$ are simple, non-negative, and form an increasing sequence with $\lambda_n \to +\infty$ as $n \to +\infty$. The corresponding unit eigenvectors $\phi_n \in L^2(0,1)$ form a Hilbert basis. The domain of the operator $\mathscr{A}$ is equivalently characterized by

$$D(\mathscr{A}) = \left\{ f \in L^2(0,1) : \sum_{n \geqslant 1} |\lambda_n|^2 |\langle f, \phi_n \rangle|^2 < +\infty \right\}.$$

Moreover we have $\mathscr{A} f = \sum_{n \geqslant 1} \lambda_n \langle f, \phi_n \rangle \phi_n$ for all $f \in D(\mathscr{A})$. For any $f \in D(\mathscr{A})$ we also define $\mathscr{A}^{1/2} f = \sum_{n \geqslant 1} \lambda_n^{1/2} \langle f, \phi_n \rangle \phi_n$.

Let $p_*, p^*, q^* \in \mathbb{R}$ be such that $0 < p_* \leqslant p(x) \leqslant p^*$ and $0 \leqslant q(x) \leqslant q^*$ for all $x \in [0,1]$, then it holds that:

$$0 \leqslant \pi^2 (n-1)^2 p_* \leqslant \lambda_n \leqslant \pi^2 n^2 p^* + q^* \tag{2.6}$$

for all $n \geqslant 1$, see e.g. (Orlov 2017). Assuming further that $q > 0$, performing an integration by parts and using the continuous embedding $H^1(0,1) \subset L^\infty(0,1)$, we obtain the existence of constants $C_1, C_2 > 0$ such that

$$C_1 \|f\|_{H^1}^2 \leqslant \sum_{n \geqslant 1} \lambda_n \langle f, \phi_n \rangle^2 = \langle \mathscr{A} f, f \rangle \leqslant C_2 \|f\|_{H^1}^2 \tag{2.7}$$

for all $f \in D(\mathscr{A})$. This implies that $f(0) = \sum_{n \geqslant 1} \langle f, \phi_n \rangle \phi_n(0)$ and $f'(0) = \sum_{n \geqslant 1} \langle f, \phi_n \rangle \phi_n'(0)$ hold for all $f \in D(\mathscr{A})$. Finally, if we further assume that $p \in \mathscr{C}^2([0,1])$, we have for any $x \in [0,1]$ that $\phi_n(x) = O(1)$ and $\phi_n'(x) = O(\sqrt{\lambda_n})$ as $n \to +\infty$, see e.g. (Orlov 2017).

2.3.2 *Homogeneous representation and spectral reduction* In order to work with an homogeneous representation of the system (2.1), we define first the change of variable:

$$w(t,x) = z(t,x) - \frac{x^2}{\cos(\theta_2) + 2\sin(\theta_2)} u(t). \tag{2.8}$$

Then introducing $v = \dot{u}$, the PDE (2.1) can be equivalently written as

$$\dot{u}(t) = v(t) \tag{2.9a}$$
$$w_t(t,x) = (p(x) w_x(t,x))_x - \tilde{q}(x) w(t,x) + a(x) u(t) + b(x) v(t) \tag{2.9b}$$
$$\cos(\theta_1) w(t,0) - \sin(\theta_1) w_x(t,0) = 0 \tag{2.9c}$$
$$\cos(\theta_2) w(t,1) + \sin(\theta_2) w_x(t,1) = 0 \tag{2.9d}$$
$$w(0,x) = w_0(x) \tag{2.9e}$$

where $a(x) = \frac{1}{\cos(\theta_2)+2\sin(\theta_2)} \{2p(x) + 2xp'(x) - x^2 \tilde{q}(x)\}$, $b(x) = -\frac{x^2}{\cos(\theta_2)+2\sin(\theta_2)}$, and $w_0(x) = z_0(x) - \frac{x^2}{\cos(\theta_2)+2\sin(\theta_2)} u(0)$.

Without loss of generality, we pick a function $q \in \mathscr{C}^0([0,1])$ and a constant $q_c \in \mathbb{R}$ such that[1]

$$\tilde{q} = q - q_c, \quad q > 0. \tag{2.10}$$

Hence, the reaction-diffusion PDE (2.9) can be rewritten in abstract form as

$$\dot{u}(t) = v(t) \tag{2.11a}$$
$$w_t(t,\cdot) = \{-\mathscr{A} + q_c \mathrm{Id}_{L^2}\} w(t,\cdot) + a u(t) + b v(t) \tag{2.11b}$$
$$w(0,\cdot) = w_0 \tag{2.11c}$$

where $\mathscr{A}$ is defined by (2.5).

We define the coefficients of projection $z_n(t) = \langle z(t,\cdot), \phi_n \rangle$, $w_n(t) = \langle w(t,\cdot), \phi_n \rangle$, $a_n = \langle a, \phi_n \rangle$, and $b_n = \langle b, \phi_n \rangle$. From (2.8) we deduce that

$$w_n(t) = z_n(t) + b_n u(t), \quad n \geqslant 1. \tag{2.12}$$

[1] If either $\theta_1 = 0$ or $\theta_2 = 0$, this condition can be relaxed to $q \geqslant 0$. This is because, in that case, (2.7) still holds true by invoking Poincaré inequality to obtain the lower bound.

Moreover, the projection of (2.11) into the Hilbert basis $(\phi_n)_{n\geqslant 1}$ reads

$$\dot{u}(t) = v(t) \tag{2.13a}$$
$$\dot{w}_n(t) = (-\lambda_n + q_c)w_n(t) + a_n u(t) + b_n v(t) \tag{2.13b}$$

with $w(t,\cdot) = \sum_{n\geqslant 1} w_n(t)\phi_n$ where the convergence of the series holds in $L^2$ norm for mild solutions and in $H^2$ norm for classical solutions. Using now (2.12) into the latter identity, the projection of (2.1) gives

$$\dot{z}_n(t) = (-\lambda_n + q_c)z_n(t) + \beta_n u(t) \tag{2.14}$$

where $\beta_n = a_n + (-\lambda_n + q_c)b_n = p(1)\{-c_{\theta_2}\phi_n'(1) + s_{\theta_2}\phi_n(1)\} = O(\sqrt{\lambda_n})$. In original coordinates we have $z(t,\cdot) = \sum_{n\geqslant 1} z_n(t)\phi_n$ with convergence of the series in $L^2$ norm.

Finally, considering classical solutions for the system trajectories, the Dirichlet measurement $y_D(t)$ given by (2.2) is expressed as the series expansion:

$$y_D(t) = z(t,0) = w(t,0) = \sum_{n\geqslant 1} w_n(t)\phi_n(0). \tag{2.15}$$

## 3. Control strategy and reduced model

### 3.1 *Control strategy*

The control strategy goes as follows. First we fix $\delta > 0$ the desired exponential decay rate for the closed-loop system trajectories. Then we define an integer $N_0 \geqslant 1$ such that $-\lambda_n + q_c < -\delta < 0$ for all $n \geqslant N_0 + 1$. We fix arbitrarily $N \geqslant N_0 + 1$, which will be specified later. The studied control strategy takes the form:

$$\hat{w}_n(t) = \hat{z}_n(t) + b_n u(t) \tag{3.1a}$$
$$\dot{\hat{z}}_n(t) = (-\lambda_n + q_c)\hat{z}_n(t) + \beta_n u(t) - l_n \left\{ \sum_{k=1}^{N} \hat{w}_k(t)\phi_k(0) - \mathrm{sat}_l(y_D(t)) \right\}, \quad 1 \leqslant n \leqslant N_0 \tag{3.1b}$$
$$\dot{\hat{z}}_n(t) = (-\lambda_n + q_c)\hat{z}_n(t) + \beta_n u(t), \quad N_0 + 1 \leqslant n \leqslant N \tag{3.1c}$$
$$u(t) = \sum_{n=1}^{N_0} k_n \hat{z}_n(t) \tag{3.1d}$$

where $l_n, k_n \in \mathbb{R}$ are the observer and feedback gains, respectively. The idea to split the dynamics of the observer in two part, a first part with correction of the error of observation (3.1b) and a second part corresponding to an open-loop estimation (3.1c), roots back to (Sakawa 1983). The former dynamics (3.1b) aims at computing the estimation $\hat{z}_n$ of the modes $z_n$ for $1 \leqslant n \leqslant N_0$ which are used to implement the control input $u$ as (3.1d). The objective of the latter dynamics (3.1c) is to improve the estimation $\hat{y}_D(t) = \sum_{k=1}^{N} \hat{w}_k(t)\phi_k(0)$ of the system output $y_D(t)$ to enhance the term of correction of the measurement error in (3.1b). In particular, the accuracy of this term improves as the dimension $N$ of the observer increases. Hence the objective is now to determine an integer $N \geqslant N_0 + 1$ so that the closed-loop system composed of (2.1) with saturated left Dirichlet boundary measurement (2.3) and the control law (3.1) is locally exponentially stable with prescribed exponential decay rate $\delta > 0$.

### 3.2 *Reduced order model for stability analysis*

We start by deriving a reduced order model for the closed-loop system composed of the PDE (2.1) and the controller (3.1). We define for $1 \leqslant n \leqslant N$ the error of observation $e_n = z_n - \hat{z}_n$ and for $N_0 + 1 \leqslant n \leqslant N$ the scaled error of observation $\tilde{e}_n = \sqrt{\lambda_n} e_n$. From the dynamics (3.1b) of observation of the $N_0$ first modes of the plant, we obtain using the deadzone nonlinearity (2.4), the series expansion (2.15), and the change of variable formula (3.1a) that

$$\begin{aligned} \dot{\hat{z}}_n(t) &= (-\lambda_n + q_c)\hat{z}_n(t) + \beta_n u(t) - l_n \left\{ \sum_{k=1}^{N} \hat{w}_k(t)\phi_k(0) - y_D(t) - \phi_l(y_D(t)) \right\} \\ &= (-\lambda_n + q_c)\hat{z}_n(t) + \beta_n u(t) + l_n \sum_{k=1}^{N_0} \phi_k(0) e_k(t) + l_n \sum_{k=N_0+1}^{N} \frac{\phi_k(0)}{\sqrt{\lambda_k}} \tilde{e}_k(t) + l_n \zeta(t) + l_n \phi_l(y_D(t)) \end{aligned} \tag{3.2}$$

for all $1 \leqslant n \leqslant N_0$ where $\zeta = \sum_{n \geqslant N+1} \phi_n(0) w_n$. We define $\tilde{z}_n = \hat{z}_n / \lambda_n$ and the vectors $\hat{Z}^{N_0} = \begin{bmatrix} \hat{z}_1 & \dots & \hat{z}_{N_0} \end{bmatrix}^\top$, $E^{N_0} = \begin{bmatrix} e_1 & \dots & e_{N_0} \end{bmatrix}^\top$, $\tilde{Z}^{N-N_0} = \begin{bmatrix} \tilde{z}_{N_0+1} & \dots & \tilde{z}_N \end{bmatrix}^\top$, and $\tilde{E}^{N-N_0} = \begin{bmatrix} \tilde{e}_{N_0+1} & \dots & \tilde{e}_N \end{bmatrix}^\top$. We also define the matrices $A_0 = \operatorname{diag}(-\lambda_1 + q_c, \dots, -\lambda_{N_0} + q_c)$, $A_1 = \operatorname{diag}(-\lambda_{N_0+1} + q_c, \dots, -\lambda_N + q_c)$, $\mathfrak{B}_0 = \begin{bmatrix} \beta_1 & \dots & \beta_{N_0} \end{bmatrix}^\top$, $\tilde{\mathfrak{B}}_1 = \begin{bmatrix} \frac{\beta_{N_0+1}}{\lambda_{N_0+1}} & \dots & \frac{\beta_N}{\lambda_N} \end{bmatrix}^\top$, $C_0 = \begin{bmatrix} \phi_1(0) & \dots & \phi_{N_0}(0) \end{bmatrix}$, $\tilde{C}_1 = \begin{bmatrix} \frac{\phi_{N_0+1}(0)}{\sqrt{\lambda_{N_0+1}}} & \dots & \frac{\phi_N(0)}{\sqrt{\lambda_N}} \end{bmatrix}$, $K = \begin{bmatrix} k_1 & \dots & k_{N_0} \end{bmatrix}$, and $L = \begin{bmatrix} l_1 & \dots & l_{N_0} \end{bmatrix}^\top$. Therefore, we obtain from (2.14), (3.1), and (3.2) that

$$u = K\hat{Z}^{N_0} \tag{3.3a}$$

$$\dot{\hat{Z}}^{N_0} = (A_0 + \mathfrak{B}_0 K)\hat{Z}^{N_0} + LC_0 E^{N_0} + L\tilde{C}_1 \tilde{E}^{N-N_0} + L\zeta + L\phi_l(y_D) \tag{3.3b}$$

$$\dot{E}^{N_0} = (A_0 - LC_0)E^{N_0} - L\tilde{C}_1 \tilde{E}^{N-N_0} - L\zeta - L\phi_l(y_D) \tag{3.3c}$$

$$\dot{\tilde{Z}}^{N-N_0} = A_1 \tilde{Z}^{N-N_0} + \tilde{\mathfrak{B}}_1 K \hat{Z}^{N_0} \tag{3.3d}$$

$$\dot{\tilde{E}}^{N-N_0} = A_1 \tilde{E}^{N-N_0}. \tag{3.3e}$$

Introducing the state vector

$$X = \operatorname{col}\left(\hat{Z}^{N_0}, E^{N_0}, \tilde{Z}^{N-N_0}, \tilde{E}^{N-N_0}\right), \tag{3.4}$$

we infer that the following truncated dynamics hold

$$\dot{X} = FX + \mathscr{L}\zeta + \mathscr{L}\phi_l(y_D) \tag{3.5}$$

where

$$F = \begin{bmatrix} A_0 + \mathfrak{B}_0 K & LC_0 & 0 & L\tilde{C}_1 \\ 0 & A_0 - LC_0 & 0 & -L\tilde{C}_1 \\ \tilde{\mathfrak{B}}_1 K & 0 & A_1 & 0 \\ 0 & 0 & 0 & A_1 \end{bmatrix}, \qquad \mathscr{L} = \begin{bmatrix} L \\ -L \\ 0 \\ 0 \end{bmatrix}. \tag{3.6}$$

REMARK 3.1 It was shown in (Lhachemi and Prieur 2021) that the pairs $(A_0, \mathfrak{B}_0)$ and $(A_0, C_0)$ both satisfy the Kalman condition. Hence we can find gains $K$ and $L$ so that $A_0 + \mathfrak{B}_0 K$ and $A_0 - LC_0$ are Hurwitz with arbitrary pole placement.

Introducing now the augmented vector $\tilde{X} = \mathrm{col}(X, \zeta, \phi_l(y_D))$ and the matrices $\tilde{K} = \begin{bmatrix} K & 0 & 0 & 0 \end{bmatrix}$ and $E = K \begin{bmatrix} A_0 + \mathfrak{B}_0 K & LC_0 & 0 & L\tilde{C}_1 & L & L \end{bmatrix}$, we obtain that the command and its first time derivative are expressed by

$$u = K\hat{Z}^{N_0} = \tilde{K}X, \qquad v = \dot{u} = K\dot{\hat{Z}}^{N_0} = E\tilde{X}. \tag{3.7}$$

Finally, in preparation of the application of Lemma 2.1 to $\phi(y_D)$, we need to express the system output $y_D$ in function of $X$ and $\zeta$. From the series expansion (2.15), the relation (2.12), the control (3.7), and recalling that $e_n = z_n - \hat{z}_n$, $\tilde{e}_n = \sqrt{\lambda_n} e_n$, and $\tilde{z}_n = \hat{z}_n / \lambda_n$, we deduce that

$$\begin{aligned} y_D &= \sum_{n \geqslant 1} \phi_n(0) w_n = \sum_{n=1}^{N} \phi_n(0) w_n + \zeta \\ &= \sum_{n=1}^{N} \phi_n(0) z_n + \sum_{n=1}^{N} \phi_n(0) b_n u + \zeta \\ &= \sum_{n=1}^{N} \phi_n(0) \hat{z}_n + \sum_{n=1}^{N} \phi_n(0) e_n + \sum_{n=1}^{N} \phi_n(0) b_n u + \zeta \\ &= \sum_{n=1}^{N_0} \phi_n(0) \hat{z}_n + \sum_{n=1}^{N_0} \phi_n(0) e_n + \sum_{n=N_0+1}^{N} \lambda_n \phi_n(0) \tilde{z}_n + \sum_{n=N_0+1}^{N} \frac{\phi_n(0)}{\sqrt{\lambda_n}} \tilde{e}_n + \sum_{n=1}^{N} \phi_n(0) b_n u + \zeta \\ &= \mathscr{C}_1 X + \mathscr{C}_2 u + \zeta \\ &= \mathscr{C} X + \zeta \end{aligned} \tag{3.8}$$

where $\mathscr{C}_1 = \begin{bmatrix} C_0 & C_0 & \lambda_{N_0+1}\phi_{N_0+1}(0) & \ldots & \lambda_N \phi_N(0) & \tilde{C}_1 \end{bmatrix}$, $\mathscr{C}_2 = \sum_{n=1}^{N} \phi_n(0) b_n$, and $\mathscr{C} = \mathscr{C}_1 + \mathscr{C}_2 \tilde{K}$.

## 4. Stability assessment

### 4.1 *Main stability result*

We are now in position to state the main result of this paper.

THEOREM 4.1 Let $\theta_1 \in (0, \pi/2]$, $\theta_2 \in [0, \pi/2]$, $p \in \mathscr{C}^2([0,1])$ with $p > 0$, $\tilde{q} \in \mathscr{C}^0([0,1])$, and $l > 0$. Let $q \in \mathscr{C}^0([0,1])$ and $q_c \in \mathbb{R}$ be such that (2.10) holds. Let $\delta > 0$ and $N_0 \geqslant 1$ be such that $-\lambda_n + q_c < -\delta$ for all $n \geqslant N_0 + 1$. Let $K \in \mathbb{R}^{1 \times N_0}$ and $L \in \mathbb{R}^{N_0}$ be such that $A_0 + \mathfrak{B}_0 K$ and $A_0 - LC_0$ are Hurwitz with eigenvalues that have a real part strictly less than $-\delta < 0$. For a given $N \geqslant N_0 + 1$, assume that there exist a symmetric positive definite $P \in \mathbb{R}^{2N \times 2N}$, positive real numbers $\alpha > 1$ and $\beta, \gamma, \mu, T, \kappa > 0$, a matrix $C \in \mathbb{R}^{1 \times 2N}$, and a real number $d \in \mathbb{R}$ such that

$$\Theta_1(\kappa) \preceq 0, \quad \Theta_2 \succeq 0, \quad \Theta_3(\kappa) \leqslant 0 \tag{4.1}$$

where

$$\Theta_1(\kappa) = \begin{bmatrix} F^\top P + PF + 2\kappa P + \alpha\gamma \|\mathscr{R}_N a\|_{L^2}^2 \tilde{K}^\top \tilde{K} & P\mathscr{L} & -TC^\top + P\mathscr{L} \\ \mathscr{L}^\top P & -\beta & -dT \\ -TC + \mathscr{L}^\top P & -dT & -2T \end{bmatrix} + \alpha\gamma \|\mathscr{R}_N b\|_{L^2}^2 E^\top E$$

$$\Theta_2 = \begin{bmatrix} P & 0 & (\mathscr{C} - C)^\top \\ 0 & \frac{\gamma}{M_\phi} & 1-d \\ \mathscr{C} - C & 1-d & \mu l^2 \end{bmatrix}$$

$$\Theta_3(\kappa) = 2\gamma \left\{ -\left(1 - \frac{1}{\alpha}\right) \lambda_{N+1} + q_c + \kappa \right\} + \beta M_\phi.$$

with $M_\phi = \sum_{n \geqslant N+1} \frac{\phi_n(0)^2}{\lambda_n} < +\infty$. Consider the block representation $P = (P_{i,j})_{1 \leqslant i,j \leqslant 4}$ with dimensions that are compatible with (3.4) and define

$$\mathscr{E} = \left\{ w \in D(\mathscr{A}) \, : \, \begin{bmatrix} \pi_{N_0} w \\ \pi_{N_0,N} \mathscr{A}^{1/2} w \end{bmatrix}^\top \begin{bmatrix} P_{2,2} & P_{2,4} \\ P_{4,2} & P_{4,4} \end{bmatrix} \begin{bmatrix} \pi_{N_0} w \\ \pi_{N_0,N} \mathscr{A}^{1/2} w \end{bmatrix} + \gamma \|\mathscr{R}_N \mathscr{A}^{1/2} w\|_{L^2}^2 < \frac{1}{\mu} \right\}. \quad (4.2)$$

Then, considering the closed-loop system composed of the plant (2.1) with saturated left Dirichlet boundary measurement (2.3) and the control law (3.1), there exists $M > 0$ such that for any initial condition $z_0 \in \mathscr{E}$ and with a zero initial condition of the observer (i.e., $\hat{z}_n(0) = 0$ for all $1 \leqslant n \leqslant N$), the system trajectory satisfies

$$\|z(t,\cdot)\|_{H^1}^2 + \sum_{n=1}^{N} \hat{z}_n(t)^2 \leqslant M e^{-2\kappa t} \|z_0\|_{H^1}^2 \quad (4.3)$$

for all $t \geqslant 0$. Moreover, for any fixed $\kappa \in (0, \delta]$, the constraints (4.1) are always feasible for $N$ selected to be large enough.

*Proof.* Consider the Lyapunov functional defined by

$$V(X,w) = X^\top P X + \gamma \sum_{n \geqslant N+1} \lambda_n \langle w, \phi_n \rangle^2, \qquad \forall X \in \mathbb{R}^{2N}, \quad \forall w \in D(\mathscr{A}). \quad (4.4)$$

The computation of the time derivative of $V$ along the system trajectories (2.13) and (3.5) gives

$$\begin{aligned} \dot{V} + 2\kappa V = {} & X^\top (F^\top P + PF + 2\kappa P) X + 2X^\top P \mathscr{L} \zeta + 2X^\top P \mathscr{L} \phi_l(y_D) \\ & + 2\gamma \sum_{n \geqslant N+1} \lambda_n \left( (-\lambda_n + q_c + \kappa) w_n + a_n u + b_n v \right) w_n. \end{aligned}$$

The use of Young's inequality gives $2\sum_{n \geqslant N+1} \lambda_n w_n a_n u \leqslant \frac{1}{\alpha} \sum_{n \geqslant N+1} \lambda_n^2 w_n^2 + \alpha \|\mathscr{R}_N a\|_{L^2}^2 u^2$ and, similarly, $2\sum_{n \geqslant N+1} \lambda_n w_n b_n v \leqslant \frac{1}{\alpha} \sum_{n \geqslant N+1} \lambda_n^2 w_n^2 + \alpha \|\mathscr{R}_N b\|_{L^2}^2 v^2$. Moreover, recalling that $\zeta = \sum_{n \geqslant N+1} \phi_n(0) w_n$, the use of Cauchy-Schwartz inequality gives $\zeta^2 \leqslant M_\phi \sum_{n \geqslant N+1} \lambda_n w_n^2$ with $M_\phi = \sum_{n \geqslant N+1} \frac{\phi_n(0)^2}{\lambda_n} < +\infty$. Hence, recalling that $\tilde{X} = \mathrm{col}(X, \zeta, \phi_l(y_D))$, we obtain that

$$\begin{aligned} \dot{V} + 2\kappa V \leqslant {} & \tilde{X}^\top \left( \begin{bmatrix} F^\top P + PF + 2\kappa P + \alpha\gamma \|\mathscr{R}_N a\|_{L^2}^2 \tilde{K}^\top \tilde{K} & P\mathscr{L} & P\mathscr{L} \\ \mathscr{L}^\top P & -\beta & 0 \\ \mathscr{L}^\top P & 0 & 0 \end{bmatrix} + \alpha\gamma \|\mathscr{R}_N b\|_{L^2}^2 E^\top E \right) \tilde{X} \\ & + \sum_{n \geqslant N+1} \lambda_n \Gamma_n w_n^2 \end{aligned}$$

where $\Gamma_n = 2\gamma\left\{-\left(1-\frac{1}{\alpha}\right)\lambda_n + q_c + \kappa\right\} + \beta M_\phi$ for all $n \geqslant N+1$. Assuming that $X \in \mathbb{R}^{2N}$ and $w \in D(\mathscr{A})$ are so that $|y_D - (CX + d\zeta)| \leqslant l$, the application of Lemma 2.1 gives $\phi_l(y_D)(\phi_l(y_D) + CX + d\zeta) \leqslant 0$. This implies that

$$\dot{V} + 2\kappa V \leqslant \tilde{X}^\top \Theta_1(\kappa)\tilde{X} + \sum_{n \geqslant N+1} \lambda_n \Gamma_n w_n^2$$

as soon as $|y_D - (CX + d\zeta)| \leqslant l$, i.e., using (3.8), $|(\mathscr{C} - C)X + (1-d)\zeta| \leqslant l$. Since $\alpha > 1$ we obtain that $\Gamma_n \leqslant \Theta_3(\kappa) \leqslant 0$ for all $n \geqslant N+1$. Moreover we recall that $\Theta_1(\kappa) \leqslant 0$. Hence, we deduce that $\dot{V} + 2\kappa V \leqslant 0$ for all $X \in \mathbb{R}^{2N}$ and $w \in D(\mathscr{A})$ so that $|(\mathscr{C} - C)X + (1-d)\zeta| \leqslant l$.

Let $X \in \mathbb{R}^{2N}$ and $w \in D(\mathscr{A})$ be such that $V(X,w) \leqslant 1/\mu$. Using Schur complement, we obtain from $\Theta_2 \succeq 0$ that

$$\frac{1}{\mu l^2}\begin{bmatrix}\mathscr{C} - C & 1-d\end{bmatrix}^\top \begin{bmatrix}\mathscr{C} - C & 1-d\end{bmatrix} \preceq \begin{bmatrix} P & 0 \\ 0 & \frac{\gamma}{M_\phi}\end{bmatrix}.$$

Recalling that $\zeta^2 \leqslant M_\phi \sum_{n \geqslant N+1} \lambda_n w_n^2$, we have

$$\frac{1}{\mu l^2}|(\mathscr{C} - C)X + (1-d)\zeta|^2 \leqslant \begin{bmatrix} X \\ \zeta \end{bmatrix}^\top \begin{bmatrix} P & 0 \\ 0 & \frac{\gamma}{M_\phi}\end{bmatrix}\begin{bmatrix} X \\ \zeta \end{bmatrix} = X^\top P X + \frac{\gamma}{M_\phi}\zeta^2 \leqslant V(X,w) \leqslant \frac{1}{\mu}, \tag{4.5}$$

which implies that $|(\mathscr{C} - C)X + (1-d)\zeta| \leqslant l$, hence $\dot{V} + 2\kappa V \leqslant 0$.

Let $w_0 \in \mathscr{E}$ be the initial condition of the PDE in homogeneous coordinates and consider zero initial conditions for the observer ($\hat{z}_n(0) = 0$ for all $1 \leqslant n \leqslant N$), giving the initial condition $z_0 = w_0$ in original coordinates. In particular one has $X(0) = \mathrm{col}(0, \pi_{N_0} z_0, 0, \pi_{N_0,N}\mathscr{A}^{1/2} z_0)$ and $V(X(0), w_0) < 1/\mu$. Using a classical contradiction argument, we infer that $V(X(t), w(t)) \leqslant 1/\mu$ for all $t \geqslant 0$, hence $\dot{V}(X(t), w(t)) + 2\kappa V(X(t), w(t)) \leqslant 0$ for all $t \geqslant 0$. This implies that $V(X(t), w(t)) \leqslant e^{-2\kappa t} V(X(0), w_0)$ for all $t \geqslant 0$. The claimed stability estimate now easily follows from the definition of $V$ and the estimates (2.7).

To complete the proof it remains to show that, for any fixed $\kappa \in (0, \delta]$, the constraints (4.1) are always feasible for $N$ selected to be large enough. We start by fixing $\alpha > 1$ arbitrarily and by setting $C = 0$ and $d = 0$. Owing to the definition (3.6) of the matrix $F$, we observe that (i) $A_0 + \mathfrak{B}_0 K + \kappa I$ and $A_0 - LC_0 + \kappa I$ are Hurwitz; (ii) $\|e^{(A_1 + \kappa I)t}\| \leqslant e^{-\kappa_0 t}$ for all $t \geqslant 0$ with the positive constant $\kappa_0 = \lambda_{N_0+1} - q_c - \kappa > 0$ independent of $N$; (iii) $\|L\tilde{C}_1\| \leqslant \|L\|\|\tilde{C}_1\|$ and $\|\tilde{\mathfrak{B}}_1 K\| \leqslant \|\tilde{\mathfrak{B}}_1\|\|K\|$ where $\|L\|$ and $\|K\|$ are independent of $N$ while $\|\tilde{C}_1\| = O(1)$ and $\|\tilde{\mathfrak{B}}_1\| = O(1)$ as $N \to +\infty$. Hence, the application of the Lemma in appendix of (Lhachemi and Prieur 2022) to the matrix $F + \kappa I$ shows that the unique solution $P \succ 0$ to the Lyapunov $F^\top P + PF + 2\kappa P = -I$ is such that $\|P\| = O(1)$ as $N \to +\infty$. We note that $\|\mathscr{L}\| = \sqrt{2}\|L\|$ and $\|\tilde{K}\| = \|K\|$ are constants independent of $N$ while $\|P\| = O(1)$ and $\|E\| = O(1)$ as $N \to +\infty$. Defining $\beta = T = N$ and $\gamma = 1/\sqrt{N}$ and recalling that we set $C = 0$ and $d = 0$, the use of Schur complement shows that we can fix $N \geqslant N_0 + 1$ large enough so that $\Theta_1(\kappa) \preceq 0$ and $\Theta_3(\kappa) \leqslant 0$. This definitely fix the order $N$ of the observer and the decision variables $P, \beta, \gamma, T > 0$. Owing to the definition of $\Theta_2$, and because $P \succ 0$ and $\gamma/M_\phi > 0$, the Schur complement implies that $\Theta_2 \succeq 0$ for $\mu > 0$ selected to be sufficiently large. This completes the proof. □

REMARK 4.1 In the context of Theorem 4.1, one can also consider non zero initial conditions $\hat{z}_n(0)$ for the observer (3.1). More precisely and under the assumptions of Theorem 4.1, there exists a constant $M > 0$ such that the estimate:

$$\|z(t,\cdot)\|_{H^1}^2 + \sum_{n=1}^{N} \hat{z}_n(t)^2 \leqslant M e^{-2\kappa t}\left\{\|z_0\|_{H^1}^2 + \sum_{n=1}^{N} \hat{z}_n(0)^2\right\}, \quad t \geqslant 0 \tag{4.6}$$

holds for all initial conditions $z_0 \in L^2(0,1)$ and $\hat{z}_n(0) \in \mathbb{R}$ selected such that, defining $w_0(x) = z_0(x) - \frac{x^2}{\cos(\theta_2)+2\sin(\theta_2)} K\hat{Z}^{N_0}(0)$, $w_0 \in D(\mathscr{A})$ and $V(X(0),w_0) = X(0)^\top PX(0) + \gamma\sum_{n\geqslant N+1}\lambda_n w_n(0)^2 < 1/\mu$. Note that the term $X(0)^\top PX(0)$ captures in particular the distances $z_n(0) - \hat{z}_n(0)$ for all $1 \leqslant n \leqslant N$. This indicates that the above stability estimate holds when the initial conditions $\hat{z}_n(0)$ of the observer are not set "too far" from the actual values $z_n(0)$ associated with the initial condition $z_0$ of the PDE.

### 4.2 *Numerical considerations*

Let the decay rate $\kappa \in (0,\delta]$ and the dimension $N \geqslant N_0+1$ of the observer be given. Constraints (4.1) of Theorem 4.1 are nonlinear w.r.t. the decision variables $P,\alpha,\beta,\gamma,\mu,T,\kappa,C,d$ due to the terms $\alpha\gamma$, $\gamma/\alpha$, $TC$, and $dT$. We discuss how LMI conditions, that remain feasible for $N$ selected large enough, can be derived from (4.1). First, we arbitrarily fix the value of $\alpha > 1$. As shown in the proof of Theorem 4.1, the obtained constraints remain feasible for a sufficiently large $N$. Now the constraints take the form of BMIs of the decision variables $P,\beta,\gamma,\mu,T,\kappa,C,d$. Second, we introduce the change of variable $\tilde{C} = TC$ and $\tilde{d} = Td$. Hence we obtain that

$$\Theta_1(\kappa) = \begin{bmatrix} F^\top P + PF + 2\kappa P + \alpha\gamma\|\mathscr{R}_N a\|_{L^2}^2 \tilde{K}^\top \tilde{K} & P\mathscr{L} & -\tilde{C}^\top + P\mathscr{L} \\ \mathscr{L}^\top P & -\beta & -\tilde{d} \\ -\tilde{C} + \mathscr{L}^\top P & -\tilde{d} & -2T \end{bmatrix}.$$

Moreover, defining $\tilde{\mu} = T^2\mu$ and

$$\tilde{\Theta}_2 = \begin{bmatrix} I & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & T \end{bmatrix}^\top \Theta_2 \begin{bmatrix} I & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & T \end{bmatrix} = \begin{bmatrix} P & 0 & (T\mathscr{C} - \tilde{C})^\top \\ 0 & \frac{\gamma}{M_\phi} & T - \tilde{d} \\ T\mathscr{C} - \tilde{C} & T - \tilde{d} & \tilde{\mu} l^2 \end{bmatrix}$$

we have $\Theta_2 \succeq 0$ if and only if $\tilde{\Theta}_2 \succeq 0$, because $T > 0$. Hence, once $\alpha > 1$ is fixed, the constraints reduce to the LMIs $\Theta_1(\kappa) \preceq 0$, $\tilde{\Theta}_2 \succeq 0$, $\Theta_3(\kappa) \leqslant 0$ with decision variables $P,\beta,\gamma,\tilde{\mu},T,\tilde{C},\tilde{d}$.

We now discuss how to optimize the decision variables in order to enlarge the estimated domain of attraction. In the framework of Theorem 4.1, let an integer $N \geqslant N_0+1$ and a $\kappa \in (0,\delta]$ such that the associated constraints (4.1) are feasible. Let a given symmetric positive definite matrix $R \in \mathbb{R}^{(N+1)\times(N+1)}$ and let $r > 0$ be selected such that

$$\mathscr{P} \triangleq \begin{bmatrix} P_{2,2} & P_{2,4} & 0 \\ P_{4,2} & P_{4,4} & 0 \\ 0 & 0 & \gamma \end{bmatrix} \preceq \frac{r}{\mu} R \tag{4.7}$$

under the constraints (4.1). Hence we deduce that

$$\begin{bmatrix} \pi_{N_0} w \\ \pi_{N_0,N}\mathscr{A}^{1/2} w \\ \|\mathscr{R}_N \mathscr{A}^{1/2} w\|_{L^2} \end{bmatrix}^\top rR \begin{bmatrix} \pi_{N_0} w \\ \pi_{N_0,N}\mathscr{A}^{1/2} w \\ \|\mathscr{R}_N \mathscr{A}^{1/2} w\|_{L^2} \end{bmatrix} \leqslant 1 \quad \Rightarrow \quad \begin{bmatrix} \pi_{N_0} w \\ \pi_{N_0,N}\mathscr{A}^{1/2} w \\ \|\mathscr{R}_N \mathscr{A}^{1/2} w\|_{L^2} \end{bmatrix}^\top \mathscr{P} \begin{bmatrix} \pi_{N_0} w \\ \pi_{N_0,N}\mathscr{A}^{1/2} w \\ \|\mathscr{R}_N \mathscr{A}^{1/2} w\|_{L^2} \end{bmatrix} \leqslant \frac{1}{\mu}$$

which shows that

$$\left\{ w \in D(\mathscr{A}) : \begin{bmatrix} \pi_{N_0} w \\ \pi_{N_0,N}\mathscr{A}^{1/2} w \\ \|\mathscr{R}_N \mathscr{A}^{1/2} w\|_{L^2} \end{bmatrix}^\top R \begin{bmatrix} \pi_{N_0} w \\ \pi_{N_0,N}\mathscr{A}^{1/2} w \\ \|\mathscr{R}_N \mathscr{A}^{1/2} w\|_{L^2} \end{bmatrix} \leqslant \frac{1}{r} \right\} \subset \mathscr{E} \tag{4.8}$$

where $\mathscr{E}$ is defined by (4.2). Hence, for given $N \geqslant N_0 + 1$ and $\kappa \in (0, \delta]$ so that the constraints (4.1) hold true, we aim at minimizing $r > 0$ under the constraints (4.1) and (4.7) with decision variables $P, \alpha, \beta, \gamma, \mu, T, C, d, r$. This minimization problem is nonlinear. In order to ease the numerical computation, it is convenient to iteratively solve the following sub-optimal problem. We start by fixing $\alpha, \mu > 0$ to values associated with a feasible solution of the constraints (4.1). Now, the only remaining nonlinearities are the products $TC$ and $dT$ in the definition of $\Theta_1(\kappa) \preceq 0$. Now one can successively fix either the value of $T$ or the values of $C, d$ to their previously computed value in order to iteratively minimize the value of $r > 0$ under LMI constraints. This approach is obviously sub-optimal but has is advantage of being numerically efficient.

## 5. Numerical illustration

We illustrate the main result of this paper, namely Theorem 4.1, with some numerical applications. We set $p = 1$, $\tilde{q} = -5$, $\theta_1 = \pi/5$, and $\theta_2 = 0$ which corresponds to a right Dirichlet boundary control. In this case, the open-loop reaction-diffusion PDE (2.1) is unstable. In this setting, the eigenstructures of the Sturm-Liouville operator, used for spectral reduction and that have been introduced in Subsection 2.3.1, are described as follows. Denoting for any integer $n \geqslant 1$ by $\mu_n$ the unique solution to $\mu_n \cot(\mu_n) = -\cot(\theta_1)$ with $\mu_n \in ((n-1)\pi, n\pi)$, we have $\lambda_n = \mu_n^2$ while the associated eigenfunctions are given by $\phi_n(x) = 2\sqrt{\frac{\mu_n}{2\mu_n - \sin(2\mu_n)}} \sin(\mu_n(1-x))$.

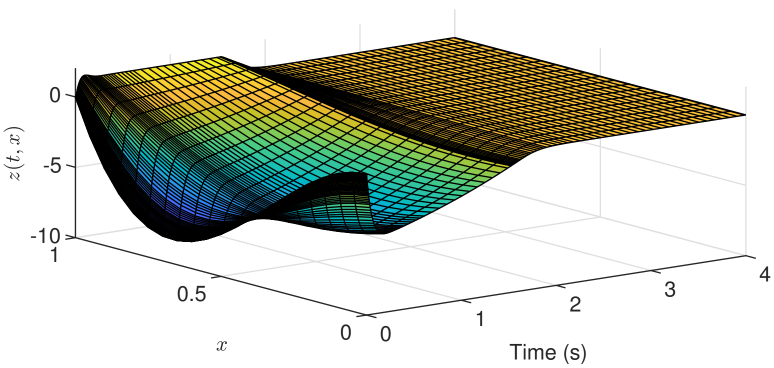
We set $N_0 = 1$, as well as the feedback and observer gains $K = -4.1481$ and $L = 10.5613$, respectively. Considering the saturation level $l = 1$ and for the exponential decay rate $\kappa = 0.2$, the constraints (4.1) are found feasible for an observer of dimension $N = 3$.

The application of the approach described in Subsection 4.2 with $P = \mathrm{diag}(I_N, 0.005)$ to enlarge the estimation of the domain of attraction of the origin gives $r = 0.0328$.

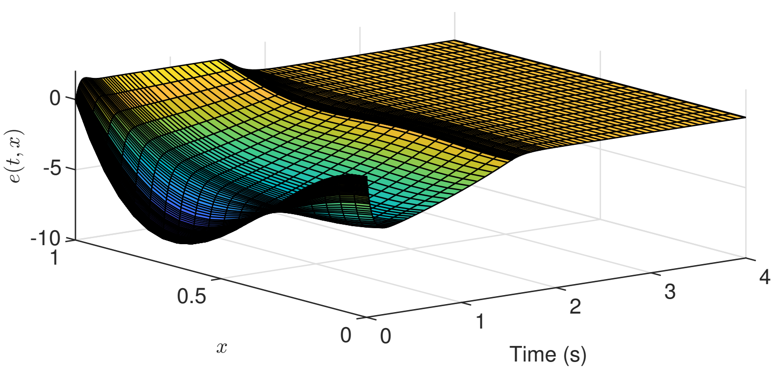
For numerical simulation, we select the initial condition of the PDE as $z_0(x) = 55.5x^2(x-1)$, for which it is checked that $z_0 \in \mathscr{E}$, along with zero initial condition of the observer. The temporal behavior of the closed-loop system is depicted in Fig. 1. It can be observed that the system trajectory converges exponentially to zero despite the saturation of the measurement shown in Fig. 1(d). This is compliant with the theoretical predictions of Theorem 4.1.

## 6. Conclusion

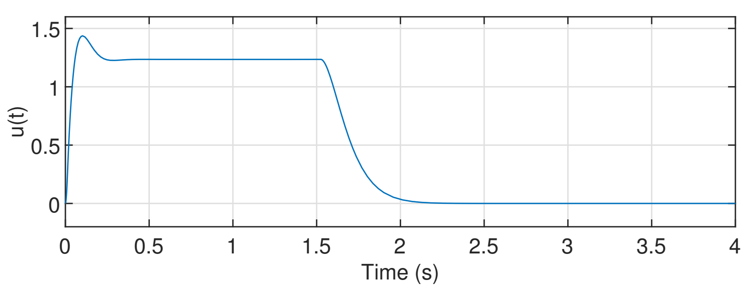
This paper has addressed the topic of output feedback stabilization of reaction-diffusion PDEs in the presence of a saturated measurement and with estimation of a subset of the domain of attraction. In contrast with the saturation of the input studied in (Mironchenko et al. 2021, Lhachemi and Prieur 2022) and for which the saturation only applies to a finite number of modes of the controller, the saturation of the measurement applies to a signal accounting for all the (infinite number and unmeasured) modes of the reaction-diffusion PDE. The controller is finite-dimensional and we showed that it always achieves the local exponential stabilization of the reaction-diffusion PDE provided the order of the observer is selected to be large enough. A distinguished feature of this approach is that the obtained stability conditions can be used to compute initial condition belonging to the domain of attraction and for which the saturation mechanism is actually active during the transient. This point has been illustrated in the numerical illustration section. Among possible extensions of this work, we can mention other optimization objectives such as the maximization of the exponential decay rate for a given domain of attraction. The possibility to include perturbations in the dynamics of the reaction-diffusion PDE in order to analyze



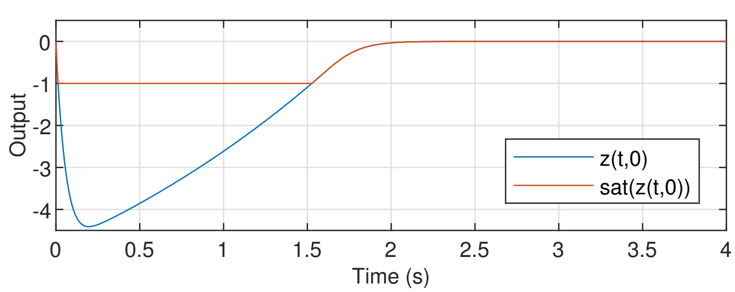

(a) State of the reaction-diffusion system $z(t,x)$




(b) Error of observation $e(t,x) = z(t,x) - \hat{z}(t,x)$




(c) Control input $u(t) = z(t,1)$




(d) Saturated Dirichlet measurement $\text{sat}_l(z(t,0))$ with $l = 1$

FIG. 1. Time evolution of the closed-loop system

their impact on the domain of attraction could also be the topic of future works.

We conclude by pointing out the main technical issue that prevents the extension of the control strategy reported in this paper to the case of a saturated Neumann measurement. More precisely, for $\theta_1 \in [0,\pi/2)$, we want to consider the left Neumann trace defined by $y_N(t) = z_x(t,0)$ for the reaction-diffusion PDE (2.1). We assume that the measurement consists in a saturated version of $y_N(t)$, i.e., $\mathrm{sat}_l(y_N(t)) = \mathrm{sat}_l(z_x(t,0))$. In that case, based on the non-saturated setting studied in (Lhachemi and Prieur 2022) with Neumann boundary measurement and the developments of this paper, a reasonable approach would be to consider the controller dynamics described by

$$\hat{w}_n(t) = \hat{z}_n(t) + b_n u(t) \tag{6.1a}$$

$$\dot{\hat{z}}_n(t) = (-\lambda_n + q_c)\hat{z}_n(t) + \beta_n u(t) - l_n \left\{ \sum_{k=1}^{N} \hat{w}_k(t)\phi_k'(0) - \mathrm{sat}_l(y_N(t)) \right\}, \quad 1 \leqslant n \leqslant N_0 \tag{6.1b}$$

$$\dot{\hat{z}}_n(t) = (-\lambda_n + q_c)\hat{z}_n(t) + \beta_n u(t), \quad N_0 + 1 \leqslant n \leqslant N \tag{6.1c}$$

$$u(t) = \sum_{n=1}^{N_0} k_n \hat{z}_n(t) \tag{6.1d}$$

where $l_n, k_n \in \mathbb{R}$ are the observer and feedback gains, respectively. However, the main technical issue here is that the residue of measurement $\zeta$ is now expressed by $\zeta = \sum_{n\geqslant N+1} \phi_n'(0) w_n$ with $\phi_n'(0) = O(\sqrt{\lambda_n})$. Hence the use of Cauchy-Schwartz inequality implies for any given $\varepsilon \in (0,1/2]$ that $\zeta^2 \leqslant M_\phi(\varepsilon) \sum_{n\geqslant N+1} \lambda_n^{3/2+\varepsilon} w_n^2$ with $M_\phi(\varepsilon) = \sum_{n\geqslant N+1} \frac{\phi_n'(0)^2}{\lambda_n^{3/2+\varepsilon}} < +\infty$. Thus we obtain a series of the type $\sum_{n\geqslant N+1} \lambda_n^\alpha w_n^2$ with $\alpha \in (3/2,2)$ compared to $\sum_{n\geqslant N+1} \lambda_n w_n^2$ in the Dirichlet measurement case. This is an issue because while the latter series can be bounded above by the Lyapunov functional $V$ defined by (4.4), which is used to obtain (4.5) and deduce the positive invariance of the ellipsoid $\mathscr{E}$, this is not the case for the former series. For this reason, the case of a saturated Neumann measurement cannot be directly handled by the method reported in this paper.

## REFERENCES

BALAS, M. J.: 1988, Finite-dimensional controllers for linear distributed parameter systems: exponential stability using residual mode filters, *Journal of Mathematical Analysis and Applications* **133**.2, 283–296.

BENZAOUIA, A., RAMI, M. A. AND EL FAIZ, S.: 2004, Stabilization of linear systems with saturation: a Sylvester equation approach, *IMA Journal of Mathematical Control and Information* **21**.3, 247–259.

BERNSTEIN, D. S. AND MICHEL, A. N.: 1995, A chronological bibliography on saturating actuators.

CAMPO, P. J. AND MORARI., M.: 1990, Robust control of processes subject to saturation nonlinearities, *Computers & Chemical Engineering* **14**.4-5, 343–358.

CORON, J.-M. AND TRÉLAT, E.: 2004, Global steady-state controllability of one-dimensional semilinear heat equations, *SIAM Journal on Control and Optimization* **43**.2, 549–569.

CORON, J.-M. AND TRÉLAT, E.: 2006, Global steady-state stabilization and controllability of 1D semilinear wave equations, *Communications in Contemporary Mathematics* **8**.4, 535–567.

CURTAIN, R.: 1982, Finite-dimensional compensator design for parabolic distributed systems with point sensors and boundary input, *IEEE Transactions on Automatic Control* **27**.1, 98–104.

DUBLJEVIC, S., EL-FARRA, N. H., MHASKAR AND P., CHRISTOFIDES, P. D.: 2006, Predictive control of parabolic PDEs with state and control constraints, *International Journal of Robust and Nonlinear Control* **16**.16, 749–772.

EL-FARRA, N. H., ARMAOU, A. AND CHRISTOFIDES, P. D.: 2003, Analysis and control of parabolic PDE systems with input constraints, *Automatica* **39**.4, 715–725.

EL HAOUSSI, F. AND TISSIR, EL H.: 2007, An LMI-based approach for robust stabilization of time delay systems containing saturating actuators, *IMA Journal of Mathematical Control and Information* **24**.3, 347–356.

GRÜNE, L. AND MEURER, T.: 2021, Finite-dimensional output stabilization of linear diffusion-reaction systems–a small-gain approach, arXiv preprint arXiv:2104.06102.

HARKORT, C. AND DEUTSCHER, J.: 2011, Finite-dimensional observer-based control of linear distributed parameter systems using cascaded output observers, *International Journal of Control* **84**.1, 107–122.

HIPPE, P.: 2006, Windup in control: its effects and their prevention, Springer Science & Business Media.

KATZ R. AND FRIDMAN, E.: 2020, Constructive method for finite-dimensional observer-based control of 1-D parabolic PDEs, *Automatica* **122**, 109285.

LASIECKA, I. AND SEIDMAN, T. I.: 2003, Strong stability of elastic control systems with dissipative saturating feedback, *Systems & Control Letters* **48**.3-4, 243–252.

LHACHEMI, H. AND PRIEUR, C.: 2021, Finite-dimensional observer-based PI regulation control of a reaction-diffusion equation, *IEEE Transactions on Automatic Control* in press.

LHACHEMI, H. AND PRIEUR, C.: 2021, Feedback stabilization of a class of diagonal infinite-dimensional systems with delay boundary control, *IEEE Transactions on Automatic Control* **66**.1, 105–120.

LHACHEMI, H. AND PRIEUR, C.: 2022, Local output feedback stabilization of a Reaction-Diffusion equation with saturated actuation, *IEEE Transactions on Automatic Control* in press.

LHACHEMI, H. AND PRIEUR, C.: 2021, Nonlinear boundary output feedback stabilization of reaction diffusion PDEs, arXiv preprint arXiv:2105.08418.

LHACHEMI, H. AND PRIEUR, C.: 2022, Finite-dimensional observer-based boundary stabilization of reaction-diffusion equations with either a Dirichlet or Neumann boundary measurement, *Automatica* **135**, 109955.

LHACHEMI, H., PRIEUR, C. AND SHORTEN, R.: 2019, An LMI condition for the robustness of constant-delay linear predictor feedback with respect to uncertain time-varying input delays, *Automatica* **109**, 108551.

LIU, Y.-Q., WANG, J.-W. AND SUN, C.-Y.: 2020, A Lyapunov-based design of dynamic feedback compensator for linear parabolic MIMO PDEs, *IMA Journal of Mathematical Control and Information* **37**.2, 455–474.

MARX, S., ANDRIEU, V. AND PRIEUR C.: 2017, Cone-bounded feedback laws for *m*-dissipative operators on Hilbert spaces, *Mathematics of Control, Signals, and Systems* **29**.4, 1–32.

MARX, S., CERPA, E., PRIEUR, C. AND ANDRIEU, V.: 2017, Global stabilization of a Korteweg–de Vries equation with saturating distributed control, *SIAM Journal on Control and Optimization* **55**.3, 1452–1480.

MIRONCHENKO, A., PRIEUR, C. AND WIRTH, F.: 2021, Local stabilization of an unstable parabolic equation via saturated controls, *IEEE Transactions on Automatic Control* **66**.5, 2162–2176.

ORLOV, Y.: 2017, On general properties of eigenvalues and eigenfunctions of a Sturm–Liouville operator: comments on "ISS with respect to boundary disturbances for 1-D parabolic PDEs", *IEEE Transactions on Automatic Control* **62**.11, 5970–5973.

PRIEUR, C., TARBOURIECH, S. AND GOMES DA SILVA JR, J. M.: 2016, Wave equation with cone-bounded control laws, *IEEE Transactions on Automatic Control* **61**.11, 3452–3463.

RUSSELL, D. L.: 1978, Controllability and stabilizability theory for linear partial differential equations: recent progress and open questions, *SIAM Review* **20**.4, 639–739.

SAKAWA, Y.: 1983, Feedback stabilization of linear diffusion systems, *SIAM Journal on Control and Optimization* **21**.5, 667–676.

SANO, H.: 2012, Stability-enhancing control of a coupled transport–diffusion system with Dirichlet actuation and Dirichlet measurement, *Journal of Mathematical Analysis and Applications* **388**.2, 1194–1204.

SLEMROD, M.: 1989, Feedback stabilization of a linear control system in Hilbert space with an a priori bounded control, *Mathematics of Control, Signals and Systems* **2**.3, 265–285.

TARBOURIECH, S., GARCIA, G., GOMES DA SILVA JR, J. M. AND QUEINNEC, I.: 2011, Stability and stabilization of linear systems with saturating actuators. Springer Science & Business Media.

WEI, Y., LU, J. AND LI, Z.: 2014, Robust stabilization of uncertain switched non-linear systems subject to actuator saturation, *IMA Journal of Mathematical Control and Information* **31**.3, 295–312.

ZACCARIAN, L. AND TEEL, A. R.: 2011, Modern anti-windup synthesis: control augmentation for actuator saturation. Princeton University Press.

ZINE, R, AND EL ALAMI, A.: 2020, Strong and weak stabilization of semi-linear parabolic systems, *IMA Journal of Mathematical Control and Information* **37**.1, 50–63.