International Journal of Control

ISSN: 0020-7179 (Print) 1366-5820 (Online) Journal homepage: https://www.tandfonline.com/loi/tcon20

Robust nonlinear state estimation for a class of infinite-dimensional systems using reduced-order models

Mouhacine Benosman & Jeff Borggaard

To cite this article: Mouhacine Benosman & Jeff Borggaard (2019): Robust nonlinear state estimation for a class of infinite-dimensional systems using reduced-order models, International Journal of Control, DOI: 10.1080/00207179.2019.1645359

To link to this article: https://doi.org/10.1080/00207179.2019.1645359

Accepted author version posted online: 16 Jul 2019.
Published online: 25 Jul 2019.

Submit your article to this journal

Article views: 14

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tcon20

INTERNATIONAL JOURNAL OF CONTROL
https://doi.org/10.1080/00207179.2019.1645359

Robust nonlinear state estimation for a class of infinite-dimensional systems using reduced-order models

Mouhacine Benosman[a] and Jeff Borggaard[b]

[a]Mitsubishi Electric Research Laboratories, Cambridge, MA, USA; [b]Department of Mathematics, Virginia Tech, Blacksburg, VA, USA

ABSTRACT
A methodology for designing robust, low-order observers for a class of spectral infinite-dimensional nonlinear systems is presented. This approach uses the low-dimensional subspace explicitly in the observer design. Then, robustness to bounded model uncertainties is incorporated using the Lyapunov reconstruction method from robust control theory. Furthermore, the proposed design includes a data-driven learning algorithm that auto-tunes the observer gains to optimise the performance of the state estimation. A numerical study using a model from fluid dynamics -Burgers equation- demonstrates the effectiveness of the proposed observer.

ARTICLE HISTORY
Received 10 April 2019
Accepted 14 July 2019

KEYWORDS
Robust observers; reduced-order models; proper orthogonal decomposition (POD); iterative feedback tuning; Burgers equation; extremum seeking

1. Introduction

The problem of designing robust observers for systems modelled by ordinary differential equations (ODEs) with parametric uncertainties and measurement noise, has been extensively studied, see e.g. Witczak, Buciakowski, Puig, Rotondo, and Nejjari (2016), Battilotti (2017) and the references therein. The extension of these results to systems modelled by PDEs (distributed parameter systems) remains a very active and challenging problem. Indeed, there are many works that utilise adaptive control to design observers for PDE systems, where both system states and parametric uncertainties are estimated, see e.g. Smyshlyaev and Krstic (2010) and references therein. However, due to the complexity of simultaneously estimating both the states and model parameters, the results are often limited to linear or semi-linear PDEs with linear parametric uncertainty. Fewer works consider passive robust control (in contrast to adaptive control) to design observers for PDEs in the presence of parametric model uncertainties and/or measurement noise. However, in the recent work (Schaum, Moreno, & Meurer, 2016), one-dimensional, semi-linear PDEs are considered and the assumption of a sector nonlinearity allows the use of dissipativity to design observers that are robust to spillover effects. In Borggaard, Gugercin, and Zietsman (2014), the authors consider the case of a PDE with a quadratic nonlinearity where the states and measurements are subject to time-varying disturbances. A MinMax approach was used to design a stabilising robust observer/controller, based on the tangent linearisation of the PDE along a steady-state solution. Then model reduction was carried out following two approaches. In one approach, an H_2-model reduction was used for the linearised system. In the second, a proper orthogonal decomposition (POD) model reduction method for nonlinear systems was used to reduce the extended Kalman filter as in Atwell, Borggaard, and King (2001). In Kharkovskaya, Efimov, Polyakov, and Richard (2018), the authors propose an interval state estimator for a class of uncertain parabolic PDE systems, under homogeneous Dirichlet boundary conditions, based on a finite-element approximation of a PDE. In Miranda, Moreno, Chairez, and Fridman (2012), a robust observer based on a super twisting algorithm, which ensures finite-time convergence, is introduced for a class of hyperbolic PDEs with bounded additive perturbations. In Feng and Guo (2017), the authors study the problem of stabilisation and observer design for the heat equation under uncertain boundary conditions. They propose a two-stage unknown input observer to estimate the uncertainty term and then observe the system states. The problem of designing a robust observer for the Boussinesq equations has been studied in Koga, Benosman, and Borggaard (2019), where the authors first used POD for model reduction, followed by a Luenberger-like observer design, based on the notion of input-state stability with respect to parameter uncertainties. These uncertainties were then estimated online using a data-driven optimisation algorithm.

In this paper, we build upon the nominal observer proposed in Balas (1981), and propose a methodology to design a robust observer for a class of spectral infinite-dimensional nonlinear systems that use a low-dimensional subspace, such as POD in the observer design. The observer is based on Lyapunov

CONTACT Mouhacine Benosman benosman@merl.com, m_benosman@ieee.org Mitsubishi Electric Research Laboratories, 201 Broadway Street, Cambridge, MA 02139, USA

© 2019 Informa UK Limited, trading as Taylor & Francis Group

reconstruction theory to 'dominate' the influence of structured model uncertainties. Furthermore, we extend this methodology so that it will auto-tune the observer gains online, using data-driven optimisation methods.

Indeed, the problem of auto-tuning feedback controllers has received much attention in the control community. It is often referred to as Iterative Feedback Tuning (IFT), and has been well-studied for the case of systems modelled by ODEs, e.g. Hjalmarsson (2002), Lequin, Gevers, Mossberg, Bosmans, and Triest (2003), Killingsworth and Krstic (2006), and Benosman (2016). However, to the best of our knowledge, IFT has not been applied in the PDE setting. In this paper, we propose the use of IFT to auto-tune the gain of a robust observer in an online setting. We follow Killingsworth and Krstic (2006), Benosman (2016), and use an extremum seeking algorithm for the tuning of the gain. This leads to the optimisation of a desired estimation performance cost function.

In the sequel, we begin by introducing some basic definitions and notation in Section 2. Section 3 is dedicated to introducing the class of nonlinear PDEs studied here, and presents the first result of the paper, namely, the nominal observer design. We use Section 4 to introduce the second result of the paper, which is the robustification of the observer under bounded model uncertainties. The third result of the paper is presented in Section 5, where we introduce the IFT version of the robust observer. Section 6 is used to present an application of the proposed robust observer and its IFT extension to a one-dimensional PDE with a quadratic nonlinearity often associated with fluid dynamics, known as Burgers equation. We conclude the paper commenting on potential future developments of this work in Section 7.

2. Basic notation and definitions

For a vector $q \in \mathbb{R}^n$, its transpose is denoted by q^T, for a matrix $C \in \mathbb{R}^{n\times m}$, the transpose is denoted by C^*. The Euclidean vector norm for $q \in \mathbb{R}^n$ is denoted by $\|\cdot\|$ so that $\|q\|_{\mathbb{R}^n} = \|q\| = \sqrt{q^T q}$. The Frobenius norm of a matrix $A \in \mathbb{R}^{n\times m}$, with elements a_{ij}, is defined as $\|A\|_F \triangleq \sqrt{\sum_{i=1}^{i=n}\sum_{j=1}^{j=m} |a_{ij}|^2}$. The Kronecker delta function is defined as: $\delta_{ij} = 0$, for $i \neq j$ and $\delta_{ii} = 1$. We shall abbreviate the time derivative by $\dot{f}(t,x) = \frac{\partial}{\partial t} f(t,x)$, and consider the following Hilbert space $\mathcal{H} = L^2(\Omega)$. We define the inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ and the associated norm $\|\cdot\|_{\mathcal{H}}$ on $\mathcal{H}$ as $\langle f,g\rangle_{\mathcal{H}} = \int_\Omega f(x)g(x)\,\mathrm{d}x$, for $f,g \in \mathcal{H}$, and $\|f\|^2_{\mathcal{H}} = \int_\Omega |f(x)|^2\,\mathrm{d}x$. A function $z(t,x)$ is in $L^2([0,t_f];\mathcal{H})$ if for each $0 \le t \le t_f$, $z(t,\cdot) \in \mathcal{H}$, and $\int_0^{t_f} \|z(t,\cdot)\|^2_{\mathcal{H}}\,\mathrm{d}t < \infty$. We will use the standard notation from distributed parameter control theory and drop the '$\cdot$' when it is understood, e.g. $z(t) = z(t,\cdot) \in \mathcal{H}$. A pseudo-inverse of an operator $\mathcal{T}$ on $\mathcal{H}$ will be denoted as $\mathcal{T}^\dagger$, and its adjoint operator on $\mathcal{H}$ is denoted by $\mathcal{T}^*$. In the sequel when we discuss the boundedness of a solution for an impulsive dynamical system, we mean uniform boundedness as defined in Haddad, Chellaboind, and Nersesov (2006, p. 67, Definition 2.12). Finally, an impulsive dynamical system is said to be well-posed, if it has well-defined distinct resetting times, admits a unique solution over a finite forward time interval, and does not exhibit any Zeno solutions, i.e. does not have an infinite number of resettings in the system over any finite time interval (Haddad et al., 2006).

3. Problem statement and observer design

We consider the state estimation problem for nonlinear systems of the form

$$\begin{aligned} \dot{z}(t) &= Az(t) + Bu(t) + h(z(t),u(t)), \quad z(0) = z_0, \\ y(t) &= Cz(t), \end{aligned} \tag{1}$$

where $z_0 \in D(A) \subset \mathcal{H}$, A is a linear operator that generates a C_0-semigroup on the Hilbert space $\mathcal{H}$, $B : \mathbb{R}^m \to \mathcal{H}$ is an input operator, $C : D(A) \to \mathbb{R}^p$ is the bounded linear operator for measurements, and h contains higher-order terms. For the well-posedness of the estimation problem, we assume that system (1) satisfies the following assumption.

Assumption 3.1: *The Cauchy problem for equation* (1) *has a solution with bounded norm $\|z(t)\|_{\mathcal{H}}$ for any initial condition $z_0 \in D(A)$, and $t > 0$.*

Furthermore, for analysis purposes we assume that h satisfies the Lipschitz-like assumption:

Assumption 3.2: *The function $h : D(A) \times \mathbb{R}^m \to [D(A)]'$ satisfies $h(0,0) = 0$ and the local Lipschitz plus constant assumption: there is a nonnegative constant β and for every pair $(z,u) \in D(A) \times \mathbb{R}^m$, there exist positive constants ϵ_z, ϵ_u, L_z, and L_u such that*

$$\|h(z,u) - h(\tilde{z},\tilde{u})\|_{\mathcal{H}} \le L_z\|z-\tilde{z}\|_{\mathcal{H}} + L_u\|u-\tilde{u}\|_{\mathbb{R}^m} + \beta,$$

for all $(\tilde{z},\tilde{u}) \in D(A) \times \mathbb{R}^m$ satisfying

$$\|z-\tilde{z}\|_{\mathcal{H}} < \epsilon_z \quad \textit{and} \quad \|u-\tilde{u}\|_{\mathbb{R}^m} < \epsilon_u.$$

We define a low-dimensional subspace $\hat{\mathcal{H}} \subset \mathcal{H}$ that inherits the norm of $\mathcal{H}$, i.e. $\|\cdot\|_{\hat{\mathcal{H}}} = \|\cdot\|_{\mathcal{H}}$, and follow the framework in Balas (1981) to design the nominal observer, while changing the roles for some operators. Consider an observer with the following structure

$$\dot{\hat{z}} = A_c\hat{z}(t) + B_c u(t) + Fy(t) + G(\hat{z}(t),u(t)), \tag{2}$$

with $\hat{z}(0) = \hat{z}_0 \in D(A_c)$, and where $A_c : \hat{\mathcal{H}} \to \hat{\mathcal{H}}$, $B_c : \mathbb{R}^m \to \hat{\mathcal{H}}$, $F : \mathbb{R}^p \to \hat{\mathcal{H}}$, and $G : \hat{\mathcal{H}} \times \mathbb{R}^m \to \hat{\mathcal{H}}$ are to be determined. Possible choices for $\hat{\mathcal{H}}$ may be the space spanned by a set of dominant eigenfunctions of A (modal approximation) or a set of basis functions obtained by performing a proper orthogonal decomposition (POD) of a collection of simulations of (1) and truncating (POD approximation), see Section 3.1.

Let $\mathcal{T} : \mathcal{H} \to \hat{\mathcal{H}}$ be the orthogonal projector from $\mathcal{H}$ to $\hat{\mathcal{H}}$ (hence, $\|\mathcal{T}\|_{\mathcal{H}} = 1$) and $\mathcal{T}^\dagger$ be the injection from $\hat{\mathcal{H}}$ into $\mathcal{H}$:

$\mathcal{T}^\dagger\hat{z} = z$ for all $\hat{z} \in \hat{\mathcal{H}} \subset \mathcal{H}$. Then we define the *reduced estimation error* as

$$e(t) = \hat{z}(t) - \mathcal{T}z(t) \in \hat{\mathcal{H}}. \tag{3}$$

This can be used as a proxy for the *state estimation error*

$$e_{\text{se}} \equiv \mathcal{T}^\dagger\hat{z} - z \in \mathcal{H}, \tag{4}$$

when $\mathcal{T}$ produces a small projection error $(z - \mathcal{T}^\dagger\mathcal{T}z)$, since

$$e_{\text{se}}(t) = \mathcal{T}^\dagger e(t) - \left(z(t) - \mathcal{T}^\dagger\mathcal{T}z(t)\right). \tag{5}$$

In fact, when $\hat{\mathcal{H}}$ is the span of r dominant POD basis functions and $\mathcal{T}_{\text{POD}}$ is the corresponding projection for a specific trajectory z, then $\mathcal{T}_{\text{POD}}$ minimises the projection error

$$\mathcal{P}(\mathcal{T}, z) = \left(\int_0^{t_f} \|z(t) - \mathcal{T}^\dagger\mathcal{T}z(t)\|_{\mathcal{H}}^2 \,\mathrm{d}t\right)^{1/2}, \tag{6}$$

over all projections $\mathcal{T}$ into subspaces of $\mathcal{H}$ with dimension r, and where t_f denotes the finite time support over which the projection error is evaluated, cf. Holmes, Lumley, and Berkooz (1998).

Remark 3.1: In practice, we can control the projection error $\mathcal{P}(\mathcal{T}, z)$ by suitable selection of the trajectory data and choosing enough basis functions r. However, we want to underline here the fact that the existence of such a basis function with clear dominant modes is only ensured for some PDEs that we denote here as *spectral PDEs*. In the case where such basis functions do not exist, e.g. hyperbolic PDEs, one could use recent results that propose more appropriate basis functions, e.g. Balajewicz, Dowell, and Noack (2013), Borggaard, Hay, and Pelletier (2007), and Rim and Mandli (2018).

Although we are free to choose B_c and G in the observer (2), to guarantee convergence we shall make the following assumptions for the remainder of this paper

$$B_c = \mathcal{T}B \quad \text{and} \quad G(\hat{z}, u) = \mathcal{T}h(\mathcal{T}^\dagger\hat{z}, u) \tag{7}$$

for all $\hat{z} \in \hat{\mathcal{H}}$ and $u \in \mathbb{R}^m$.

We can now state our first result.

Theorem 3.1: *Consider the system described by* (1) *under Assumptions* 3.1, 3.2, *for which we associate the state observer defined by* (2) *and* (7). *We assume that F, A_c, and $\mathcal{T}$ satisfy the conditions*

$$[A_c\mathcal{T} - \mathcal{T}A + FC]\,z = 0, \quad \textit{for all } z \in D(A), \tag{D0}$$

$$\|\exp(A_c t)\|_{\hat{\mathcal{H}}} \leq M\exp(-\delta t), \quad \textit{for all } t > 0 \tag{D1}$$

and,

$$\delta > ML_z, \tag{D2}$$

where $M \geq 1$ and $\delta > 0$. Then we can guarantee the exponential stability of the estimation error, $e(t)$ in (3). *Namely, there exists a constant c, depending on δ, M, the initial error $\|e(0)\|_{\hat{\mathcal{H}}}$, and the $\mathcal{P}(\mathcal{T}, z)$ in* (6) *such that*

$$\|e(t)\|_{\hat{\mathcal{H}}} \leq c\,\exp((ML_z - \delta)t)\|e(0)\|_{\hat{\mathcal{H}}}, \tag{8}$$

where,

$$c = M\left\{\|e(0)\|_{\hat{\mathcal{H}}} + L_z\left(\frac{\exp(2\delta t_f) - 1}{2\delta}\right)^{1/2}\Pi(\mathcal{P}(\mathcal{T}, z), \beta)\right\}, \tag{9}$$

and

$$\Pi(\mathcal{P}(\mathcal{T}, z), \beta) = \left(\int_0^{t_f}\left(\|z(t) - \mathcal{T}^\dagger\mathcal{T}z(t)\|_{\mathcal{H}} + \frac{\beta}{L_z}\right)^2 \mathrm{d}t\right)^{1/2}. \tag{10}$$

Proof: If we differentiate (3) with respect to time and use (1) and (2), we find

$$\begin{aligned}
\dot{e}(t) &= \dot{\hat{z}}(t) - \mathcal{T}\dot{z}(t) \\
&= A_c\hat{z}(t) + B_c u(t) + Fy(t) + G(\hat{z}(t), u(t)) \\
&\quad - \mathcal{T}\left[Az(t) + Bu(t) + h(z(t), u(t))\right] \\
&= A_c e(t) + \left[A_c\mathcal{T} - \mathcal{T}A + FC\right] z(t) \\
&\quad + \left[B_c - \mathcal{T}B\right] u(t) + N(e(t), z(t), u(t)),
\end{aligned} \tag{11}$$

where $N(e, z, u) \equiv G(e + \mathcal{T}z, u) - \mathcal{T}h(z, u)$. The second term on the right hand side vanishes if we require condition (D0) and the third vanishes using our choice of B_c in (7). Thus, we are left with

$$\dot{e}(t) = A_c e(t) + N(e(t), z(t), u(t)), \tag{12}$$

or

$$e(t) = \exp(A_c t)e(0) + \int_0^t \exp(A_c(t - s))N(e(s), z(s), u(s))\,\mathrm{d}s. \tag{13}$$

The matrix A_c is stable from (D1). Thus, we will exploit our choice of G in (7) and the local Lipschitz plus bounded condition (Assumption 3.2) on h to bound the integral term. First of all,

$$\begin{aligned}
\|N(e, z, u)\|_{\mathcal{H}} &= \|\mathcal{T}h(\mathcal{T}^\dagger(e + \mathcal{T}z), u) - \mathcal{T}h(z, u)\|_{\mathcal{H}} \\
&\leq L_z\|\mathcal{T}^\dagger e + \mathcal{T}^\dagger\mathcal{T}z - z\|_{\mathcal{H}} + \beta \\
&\leq L_z\left(\|e\|_{\hat{\mathcal{H}}} + \|\mathcal{T}^\dagger\mathcal{T}z - z\|_{\mathcal{H}} + \frac{\beta}{L_z}\right).
\end{aligned} \tag{14}$$

Therefore, (13) leads to

$$\begin{aligned}
\|e(t)\|_{\hat{\mathcal{H}}} &\leq M\exp(-\delta t)\|e(0)\|_{\hat{\mathcal{H}}} \\
&\quad + \int_0^t M\exp(-\delta(t - s))L_z\|e(s)\|_{\hat{\mathcal{H}}}\,\mathrm{d}s \\
&\quad + ML_z\exp(-\delta t)\int_0^t \exp(\delta s)\left(\|\mathcal{T}^\dagger\mathcal{T}z(s)\right. \\
&\quad \left. - z(s)\|_{\mathcal{H}} + \frac{\beta}{L_z}\right)\mathrm{d}s.
\end{aligned}$$

By applying the Cauchy-Schwarz inequality to the last term above and using the Gronwall-Reid inequality, we obtain

$$\|e(t)\|_{\hat{\mathcal{H}}} \leq c \exp((ML_z - \delta)t)\|e(0)\|_{\hat{\mathcal{H}}}, \tag{15}$$

where c is given in (9). Finally, using assumption (D2) in equation (8) gives us exponential stability of the error. ■

Remark 3.2: Condition (D0) can be exactly satisfied for a class of bounded linear operators $\mathcal{T}$, as proven in (Theorem 3.2, Balas, 1981). However, in the more practical context of POD-based realisation of the observer, presented here in Section 3.1, we will approximate condition (D0), such that the residual effect of its approximation does not change the exponential convergence result of Theorem 3.1, see Remark 3.6.

Remark 3.3: The influence of the projection error $\mathcal{P}(\mathcal{T}, z)$ on the reduced estimation error $e(t)$ appears explicitly in the calculation of the constant c above. Indeed, this is one advantage of the estimator derived above and explicitly links the ROM-based estimation error and the projection error. Many reduce-then-design approaches to design observers for PDE systems, e.g. Koga et al. (2019), first build a reduced-order model (ROM) by projection, then separately build an observer for the ROM. The separation of the projection subspaces $\hat{\mathcal{H}}$ from the observer design in the reduce-then-design approaches miss the explicit connection that we have included by using $\mathcal{T}$ in assumption (D0) as well as in the reduced nonlinear operator (7), which ultimately leads to the definition of c in (9). Another point that further differentiates our approach from the reduce-then-design methods, is that the later methods when applied to some type of PDEs can lead to an unstable reduced order model (ROM). This ROM then needs to be stabilised first before designing a ROM-based observer, e.g. Benosman, Borgaard, San, and Kramer (2017) and Koga et al. (2019). In this work, we do not have to impose any stability constraints on the projection $\mathcal{T}A$, we only require that it satisfies condition (D0). Finally, we can also underline that contrary to the classical ROM-based Luenberger-like observer design, e.g. Koga et al. (2019), the proposed observer (2) does not explicitly use an output-error injection term in its design.

Remark 3.4: The upper bound in (15) shows an exponential decrease of the estimation error norm, however, this bound can be large in the case of large values of β, since c in (9) is directly proportional to β. We will see in Section 4 that this upper-bound estimate can be improved by a robustification of the observer, in the case of bounded additive model uncertainties.

3.1 Observer design based on the proper orthogonal decomposition

We first compute the proper orthogonal decomposition (POD) from solutions to (1) then use this as a basis for $\hat{\mathcal{H}}$. Since POD with Galerkin projection is a well-known model reduction method for nonlinear problems, we will keep this discussion brief and refer the interested reader to Holmes et al. (1998) and Kunisch and Volkwein (2007).

Given a trajectory (or snapshots) of (1)

$$\mathcal{S} = \{z(t,\cdot) \in \mathcal{H} \mid t \in [0, t_f]\}, \tag{16}$$

the spatial autocorrelation function K is defined as $K(x,\bar{x}) = \frac{1}{t_f}\int_0^{t_f} z(t,x)z^*(t,\bar{x})\,\mathrm{d}t$, and is well defined when $z(t,x)$ is in $L^2([0,t_f];\mathcal{H})$. The function K is used as the kernel of the Fredholm problem $\int_\Omega K(x,\bar{x})\phi(\bar{x})\,\mathrm{d}\bar{x} = \lambda\phi(x)$. Using Fredholm theory, there exist solution pairs $\{(\lambda_i,\phi_i)\}_{i=1}^{\infty}$, where the *POD eigenvalues* $\{\lambda_i\}_{i=1}^{\infty}$ satisfy $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ with the only accumulation point at 0, and the *POD basis functions* $\{\phi_i\}_{i=1}^{\infty}$ are orthonormal functions, $\langle\phi_i,\phi_j\rangle_{\mathcal{H}} = \delta_{ij}$. We now consider the reduced basis of the first r terms based on a desired projection error (6): $\hat{\mathcal{H}}_r = \mathrm{span}\{\phi_1(\cdot),\phi_2(\cdot),\ldots,\phi_r(\cdot)\}$, and approximate solutions to (1) in $\hat{\mathcal{H}}_r$ using

$$z_r^{pod}(t,\cdot) = \sum_{i=1}^{r} q_i(t)\phi_i(\cdot) \;\in \hat{\mathcal{H}}_r, \tag{17}$$

where q_i, $i = 1,\ldots,r$ are the POD projection coefficients.

We then define the (orthogonal) projection operator $\mathcal{T} \equiv \mathcal{T}_{\mathrm{POD}} : \mathcal{H} \to \hat{\mathcal{H}}_r$ as follows

$$[\mathcal{T}_{\mathrm{POD}} z](\cdot) = \sum_{i=1}^{r} \phi_i(\cdot)\langle\phi_i, z\rangle_{\mathcal{H}}. \tag{18}$$

The pseudo-inverse of $\mathcal{T}$ is the injection of $\hat{\mathcal{H}}_r$ into $\mathcal{H}$. Thus $\mathcal{T}^\dagger\hat{z} = z$ for all $\hat{z} \in \hat{\mathcal{H}}_r$ and since $\mathcal{T}$ is a projection operator, we have $\mathcal{T}\mathcal{T}^\dagger = \mathcal{I}_r$.

Next, we define $A_c : \hat{\mathcal{H}}_r \to \hat{\mathcal{H}}_r$ as

$$A_c = \mathcal{T}^{\dagger *} A \mathcal{T}^\dagger. \tag{19}$$

With this selection, we can show that for any $\hat{z} \in \hat{\mathcal{H}}_r$ with $\|\hat{z}\|_{\hat{\mathcal{H}}} = 1$, the following holds: $\langle A_c\hat{z},\hat{z}\rangle = \langle A\mathcal{T}^\dagger\hat{z}, \mathcal{T}^\dagger\hat{z}\rangle \leq \max_{\|z\|_{\mathcal{H}}=1}\langle Az, z\rangle$.

Remark 3.5: If A is self-adjoint and exponentially stable, the suggested choice for A_c in (19), ensures that (D1) is satisfied, e.g. see (Definition 7, Jacobson & Nett, 1988). Condition (D2) may naturally be enforced with our choice of the projection operator $\mathcal{T}$ and the local Lipschitz constant associated with the solution we are estimating. However, one may need to modify the construction of A_c to simultaneously ensure exponential stability, as well as, impose a sufficient decay constant for A_c, cf. Benosman et al. (2017), Noack et al. (2008), and Wang, Akhtar, Borggaard, and Iliescu (2012). For example, by substituting $\tilde{A}_c = A_c + \hat{A}_c$ for A_c, where $\hat{A}_c$ is used to tune the decay rate of the new $\tilde{A}_c$ matrix.

Condition (D0) is the most challenging to satisfy. We define F as

$$F = (\mathcal{T}A - A_c\mathcal{T})C^\dagger, \tag{20}$$

where $C^\dagger$ is a left pseudo-inverse of the bounded linear operator C, e.g. Beutler (1965).

Remark 3.6: We want to underline here that in applications, and due to the finite number of sensors (even sparse in most real-life applications), it is clear that equation (20), which stems from our POD formulation of the observer, constitutes an approximation in a least-squares sense of the exact condition (D0) This is due to the fact that the pseudo-inverse $C^{\dagger}$ is only an approximation of the exact left-inverse of C, e.g. Beutler (1965, pp. 451–452). This approximation could also be obtained by directly minimising the term $[A_c\mathcal{T} - \mathcal{T}A + FC]z$ for $z \in \text{span}\{\phi_i\}$, i.e. along a simulated solution of the system. Another solution would be to use the matrices decomposition used in Witczak et al. (2016) for solving a similar Sylvester equation (in the ODE setting). However, such solution will also be an approximation in our case of a non-square measurement operator C, i.e. less sensors than the large state variables number obtained from discretisation. In essence, what we need is for the term $[A_c\mathcal{T} - \mathcal{T}A + FC]z(t)$ to be as small as achievable, under the constraint of finite number of sensors. Indeed, the fact that condition (D0) is not exactly satisfied does not change the exponential convergence of the error shown in Theorem 3.1, since if we denote by $\text{res}_{Sylvester}$ the residual error in solving the Sylvester equation $A_c\mathcal{T} - \mathcal{T}A + FC = 0$, using (20), then due to Assumption 3.1, one can bound the norm of the residual term $\text{res}_{Sylvester}z$, which can then be included in the constant term β when computing the upper-bound of N in (14). Additionally, the effect of this bounded residual term can be compensated for by the robustification of the observer, as presented in the next section.

4. Robustification of the observer

In this section we use tools from robust control theory, i.e. Lyapunov redesign techniques, e.g. Khalil (1996) and Benosman and Lum (2010), to robustify the nominal observer developed in the previous section. Let us consider the case where the system (1) contains an uncertainty on h, as follows

$$\dot{z}(t) = Az(t) + Bu(t) + h(z(t), u(t)) + \Delta h(z(t)), \qquad (21a)$$

$$y(t) = Cz(t), \qquad (21b)$$

from $z(0) = z_0$, where the uncertainty $\Delta h : \mathcal{H} \to \mathcal{H}$, satisfies the following assumption.

Assumption 4.1: *The uncertainty $\Delta h : \mathcal{H} \to \mathcal{H}$, is uniformly bounded: there exists a constant $\Delta h_{\max} > 0$ such that $\|\Delta h(z)\|_{\mathcal{H}} \leq \Delta h_{\max}$, $\forall\, z \in \mathcal{H}$.*

Now, if we examine the dynamics of the observer (2), we see that the observer convergence relies on the design of the nonlinear function G, in (7). To robustify the nominal design presented in Section 3, and account for the additional uncertainty term Δh, we use a Lyapunov redesign approach and add an additional term to G. The robust observer is now written as

$$\dot{\hat{z}}(t) = A_c\hat{z}(t) + B_c u(t) + Fy(t) + G(\hat{z}, u) + \Delta G(\hat{z}), \qquad (22)$$

with A_c, B_c, F, G satisfying conditions (7), (D0), (D1), (D2), and where $\Delta G : \hat{\mathcal{H}} \to \hat{\mathcal{H}}$, must be designed to compensate for any negative impact that the uncertainty Δh might have on the exponential stability of e obtained in (8). Carrying out a similar analysis for the robust observer (22), under (7), and (D0), the associated error dynamics satisfy

$$\begin{aligned} \dot{e}(t) &= A_c e(t) + G\,(e(t) + Tz(t), u(t)) \\ &\quad - \mathcal{T}h(z(t), u(t)) + \Delta G(\hat{z}) - \mathcal{T}\Delta h(z). \end{aligned} \qquad (23)$$

In the remainder of this section, we will try to recover at least the convergence of e to a positively invariant set with a radius that we can control, regardless of the form of the bounded uncertainty Δh. We summarise the first result of this section in the following theorem.

Theorem 4.1: *Consider the error dynamics (23) for the observer (22) and (7), tracking the uncertain system (21). Let h and Δh satisfy Assumptions 3.2 and 4.1, respectively. Define the compensation term ΔG as*

$$\Delta G(\hat{z}) = k\Delta h_{\max}\tilde{C}^*\tilde{C}e, \qquad (24)$$

for $k < 0$, and $\tilde{C}$ satisfying

$$\tilde{C}\mathcal{T} = C. \qquad (25)$$

Then under Assumption 3.1, and conditions (D0), (D1), and (D2), the solution of the error dynamics equation (23) converges to the invariant set

$$S = \{e \in \hat{\mathcal{H}},\ satisfying,\ k\|e\|_{\hat{\mathcal{H}}}\lambda_{\min}(\tilde{C}^*\tilde{C}) + 1 \geq 0\},$$

and the estimation error upper-bound is given by

$$\begin{aligned} \|e(t)\|_{\hat{\mathcal{H}}} \leq \frac{-1}{k\lambda_{\min}(\tilde{C}^*\tilde{C})} + \left(\|e(0)\|_{\hat{\mathcal{H}}} + \frac{1}{k\lambda_{\min}(\tilde{C}^*\tilde{C})}\right) \\ \exp(k\Delta h_{\max}\lambda_{\min}(\tilde{C}^*\tilde{C})t). \end{aligned} \qquad (26)$$

Proof: We define the Lyapunov function as

$$V(e) = \tfrac{1}{2}\langle e, e\rangle_{\hat{\mathcal{H}}}, \qquad (27)$$

then show that our design for ΔG in (24) compensates for the uncertainty Δh in (21), by providing an asymptotic decrease in V as the system evolves.

Taking the derivative along solutions leads to:

$$\begin{aligned} \dot{V}(e(t)) &= \langle e(t), A_c e(t) + G\left(\hat{z}(t), u(t)\right) - \mathcal{T}h(z(t), u(t))\rangle \\ &\quad + \langle e(t), \Delta G(\hat{z}(t)) - \mathcal{T}\Delta h(z(t))\rangle. \end{aligned}$$

Note that, due to the exponential stability of e in the nominal case (when $\Delta G \equiv 0$, $\Delta h \equiv 0$), the fact that G is given by (7), and using the conditions (D0), (D1), and (D2), we know that $\dot{V}$ is negative along the solution to the nominal error dynamics (12). Thus, the first term of the right-hand-side is negative, and we can write

$$\begin{aligned} \dot{V}(e(t)) &\leq \langle e(t), \Delta G(\hat{z}(t)) - \mathcal{T}\Delta h(z(t))\rangle, \\ &\leq \langle e(t), \Delta G(\hat{z}(t))\rangle + \|e(t)\|_{\hat{\mathcal{H}}}\Delta h_{\max}. \end{aligned}$$

Now to compensate for the effect of the Δh term, and preserve the decrease of V along the new error dynamics (23), we define the term ΔG as

$$\Delta G(\hat{z}) = k\Delta h_{\max}\tilde{C}^*\tilde{C}e, \quad k < 0. \tag{28}$$

This allows us to bound $\dot{V}$ as

$$\begin{aligned}\dot{V} &\leq k\|e\|^2_{\hat{\mathcal{H}}}\Delta h_{\max}\lambda_{\min}(\tilde{C}^*\tilde{C}) + \|e\|_{\hat{\mathcal{H}}}\Delta h_{\max} \\ &\leq (k\|e\|_{\hat{\mathcal{H}}}\lambda_{\min}(\tilde{C}^*\tilde{C}) + 1)\Delta h_{\max}\|e\|_{\hat{\mathcal{H}}}.\end{aligned}$$

This proves convergence of the error to the invariant set, e.g. see Khalil (1996): $S = \{e \in \hat{\mathcal{H}},\ \text{satisfying } k\|e\|_{\hat{\mathcal{H}}}\lambda_{\min}(\tilde{C}^*\tilde{C}) + 1 \geq 0\}$.

Finally, to establish the upper-bound for $\|e\|_{\hat{\mathcal{H}}}$, we use the following classical argument: We define $Y = \|e\|_{\hat{\mathcal{H}}}$, which leads to

$$\dot{V} = Y\dot{Y}, \tag{29}$$

we can then write the inequalities

$$\dot{Y} \leq \Delta h_{\max}(Yk\lambda_{\min}(\tilde{C}^*\tilde{C}) + 1). \tag{30}$$

Thus, Y is bounded by the solution of the ordinary differential equation

$$\dot{y} = \Delta h_{\max}(yk\lambda_{\min}(\tilde{C}^*\tilde{C}) + 1), \quad y(0) = Y(0), \tag{31}$$

which finally allows us to write the inequality (26). ∎

Remark 4.1: The introduction of the operator $\tilde{C}$ in the definition of ΔG in (24) is not required to show stabilisation of the estimation error to the invariant set S. Indeed, the upper-bound on $\dot{V}$ can be made negative without the need of $\tilde{C}$. However, to make the observer implementable, one cannot consider cases where the full state z is available for feedback. Hence the need to project z into the space of measurements through the use of the mapping $\tilde{C}$. By further defining $\tilde{C}$ to satisfy $\tilde{C}\mathcal{T} = C$, we can implement the robust portion of the observer as follows:

$$\begin{aligned}\Delta G &= k\Delta h_{\max}\tilde{C}^*\tilde{C}e, \\ &= k\Delta h_{\max}\tilde{C}^*\tilde{C}(\hat{z} - \mathcal{T}z), \\ &= k\Delta h_{\max}\tilde{C}^*(\tilde{C}\hat{z} - Cz), \\ &= k\Delta h_{\max}\tilde{C}^*(\tilde{C}\hat{z} - y),\end{aligned} \tag{32}$$

which only requires the observer states $\hat{z}$, and the measured output y.

Remark 4.2: The robustification of the observer allows us to obtain a tighter upper-bound of the estimation error norm given by (26), since it is inversely proportional to the observer gain k, which can be selected high enough to tighten this upper-bound.

The passive robustification presented above guarantees asymptotic convergence of the observer. However, this robustness might lead to poor transient performance in practice. Thus, one is also interested in improving the transient performance of the observer. For this reason, we want to improve the passive robust observer presented in this section by complementing it with an active learning step. This step learns the best (in an optimal sense that we define later) observer feedback gain k.

5. Learning-based tuning of the observer

In this section we want to merge together the passive robust observer given by (22), and (24), with an active learning algorithm, to improve the performance of the observer. Indeed, one parameter that could benefit from online tuning is the robust observer gain k defined in (24). If we examine the results of Theorem 4.1, we see that the estimation error upper-bound (invariant set radius) decreases with the decrease of the negative feedback gain. However, if we are concerned with more than asymptotic convergence to an invariant set, we need to tune the feedback gain k to achieve other objectives. For instance, if one is interested in optimising the transient behaviour of the observer, the gain k needs to be tuned to optimise a transient estimation cost performance. To find the optimal value of the observer gain, we propose to use a data-driven optimisation algorithm to auto-tune the gain online, while the observer is estimating the system states. This problem is strongly related to iterative feedback tuning (IFT), e.g. Hjalmarsson (1998), Lequin et al. (2003), Hjalmarsson (2002), Benosman (2016), and Killingsworth and Krstic (2006). We will follow Killingsworth and Krstic (2006), Benosman (2016), and use an extremum seeking (ES)-based auto-tuning approach. We first write the feedback gain as

$$k = k_{\text{nom}} + \delta k, \quad k_{\text{nom}} < 0, \tag{33}$$

where k_{nom} represents the nominal value of the observer gain, and δk is the necessary adjustment of the gain to improve the transient performance of the observer. We then define the learning cost function

$$\begin{aligned}Q(\delta k) &= \int_0^T ||e_y||^2_{\hat{\mathcal{H}}}\,\mathrm{d}t, \\ e_y(\delta k) &= \hat{y}(t; \delta k) - y(t), \\ \hat{y} &= C\hat{z},\end{aligned} \tag{34}$$

where $T > 0$, $\hat{z}$ is solution of the observer (22), (24), and y is the measured output. Furthermore, for analysis purposes, we will need the following assumptions on Q.

Assumption 5.1: *The cost function $Q(\delta k)$ in (34) has a local minimum at $\delta k = \delta k_*$.*

We propose to use the following time-varying amplitude-based ES algorithm, introduced in Tan, Nesic, Mareels, and Astolfi (2009), to tune δk

$$\begin{aligned}\dot{x}_k &= -\delta_k\omega_k \sin(\omega_k t)Q(\delta k), \\ \delta k(t) &= x_k(t) + a_k \sin(\omega_k t), \\ \dot{a}_k &= -\delta_k\omega_k\epsilon_k a_k,\end{aligned} \tag{35}$$

where $\delta_k > 0, \omega_k > 0, \epsilon_k > 0$. We summarise the gain auto-tuning algorithm in the following theorem.

Theorem 5.1: *Consider the observer (7), (22), and (24), where the gain k is tuned iteratively, with each iteration being of finite time T, such that the state is reset over the tuning iteration*

$j = 1, 2, \ldots,$ as $\hat{z}(jT) = \hat{z}_0,\ j = \{1, 2, \ldots\}$, and the gain–over iterations–is defined as

$$k(t) = k_{\text{nom}} + \Delta k(t), \quad k_{\text{nom}} > 0$$
$$\Delta k(t) = \delta k((j-1)T), \quad (j-1)T \le t < jT,\ j = 1, 2, 3 \ldots \tag{36}$$

where δk is defined by the forward first order Euler discretisation of (34), (35), *with a time step equal to T. Then, the impulsive dynamic* (22), (24), (34), (35), *and* (36), *is well-posed, the state vector $\hat{z}$ is uniformly bounded, and under Assumption* 5.1, *the gain k converges to a neighbourhood of its local optimum value $k_{\text{nom}} + \delta k_*$.*

Proof: The proof follows similar arguments as the one used in proving Theorem 2 of Benosman (2016). Indeed, we first observe that the closed-loop system (7), (22), (24), (36), (34), and (35) can be viewed as an impulsive time-dependent dynamical system, Haddad et al. (2006, pp. 18–19), with the trivial resetting law $\Delta\hat{z}(t) = \hat{z}_0$, for $t = jT,\ j \in \{1, 2, \ldots\}$. In this case the resetting times given by $jT,\ T > 0\ j \in \{1, 2, \ldots\}$, are well defined and distinct. Furthermore, due to Assumption 3.2 and the smoothness of (7), (22), and (24) (within each learning iteration), this impulsive dynamic system admits a unique solution in forward time, for any initial condition $\hat{z}_0 \in \hat{\mathcal{H}}$ (Haddad et al., 2006, p. 12). Finally, the fact that $T \neq 0$ excludes a Zeno behaviour over a finite time interval (only a finite number of resets are possible over a finite time interval). Next, if we consider the error dynamic (23) with the initial error $e_0 = \hat{z}(0) - \mathcal{T}z(0)$, then under the conditions of Theorem 4.1, there exists, for any given time-interval $(j-1)T \le t < jT$, for any given $j \in \{1, 2, \ldots\}$, a Lyapunov function $V_j = \frac{1}{2}\langle e, e\rangle$, such that, $\dot{V}_j \le (k_j\|e\|_{\hat{\mathcal{H}}}\lambda_{\min}(\tilde{C}^*\tilde{C}) + 1)\Delta h_{\max}\|e\|$, where k_j is the gain for iteration $j \in \{1, 2, \ldots\}$. This shows that e, starting from e_0 (for all the iterations $j \in \{1, 2, \ldots\}$) is steered $\forall t \in [(j-1)T,\ jT[$, towards the invariant set $S_j = \{e \in \hat{\mathcal{H}},\ \text{s.t.},\ k_j\|e\|_{\hat{\mathcal{H}}}\lambda_{\min}(\tilde{C}^*\tilde{C}) + 1 \ge 0\}$. Furthermore, since at each switching point, i.e. each new iteration j, we reset the system from the same bounded initial condition e_0, we can conclude uniform boundedness of the tracking error e. Next, since we restart each learning iteration from the same inial condition e_0, then the cost function (34) is well defined as a function of the optimisation parameter δk. Finally, by Theorem 1, in Tan et al. (2009) and accounting for the global $o(T)$ error of a first-order forward Euler discretisation, we can conclude, under Assumption 5.1, the convergence of the extremum seeker (35) to a neighbourhood $o(T)$ of the local optimal value δk_*. ■

Remark 5.1: We decided to use the ES algorithm of Tan et al. (2009) for two reasons: (1) Under stronger assumptions, i.e. existence and uniqueness of a global minimum of Q (Assumption 3, in Tan et al., 2009), and another technical assumption on the equilibrium solutions of the averaged system of the ES dynamics (Assumption 4, in Tan et al., 2009), one can claim semi-global convergence to a neighbourhood of the global minimum, i.e. semi-global practical stability of the global minimum (Theorem 1, in Tan et al., 2009), even in the case of existence of minima. (2) Due to the asymptotic decrease of the dither amplitude, $a_k(t)$, which is a solution of the stable dynamics given by the third equation in (35), the ES algorithm converges to a tight neighbourhood of the minimum (local or global), with less residual dither oscillations compared to other classical dither-based ES algorithms with constant dither signal amplitude, e.g. Tan, Nesic, and Mareels (2006) and Krstic (2000). The latter point can be easily seen from the second equation in (35), where one observes that the oscillations in δk introduced by the dither signal, vanishes with $a_k(t)$. However, we want to emphasise that in the absence of these assumptions, the algorithm still ensures local convergence to a local extremum, which means the auto-tuning will still have a beneficial effect on the observer performance.

Remark 5.2: Theorem 5.1 does not directly deal with the convergence of the observer, but it deals with the optimisation of the transient solution of the observer. Indeed, in Theorem 5.1, we analyse the convergence of the auto-tuning algorithm ((35), (34), and (35)) that is introduced to auto-tune the gain $k < 0$ of the observer. In other words, instead of tuning the negative gain k manually, where each optimal value would depend on the new initial conditions and optimises its own transient tracking performance defined by the cost Q in (34), we use an auto-tuning optimisation algorithm that will tune the gain online, and automatically find an optimal gain from the set of all stabilising gains. This idea is usually used in gain tuning of feedback controls, and is referred to as iterative feedback tuning (IFT), e.g. Benosman (2016). We use it here as gain tuning for our observer.

6. An application example from fluid dynamics: the 1D Burgers equation

We consider estimating solutions to the 1D damped Burgers equation, e.g. Burns and Kang (1990)

$$\frac{\partial z(t,x)}{\partial t} + z(t,x)\frac{\partial z(t,x)}{\partial x} = \mu\frac{\partial^2 z(t,x)}{\partial x^2} - \gamma z(t,x), \tag{37}$$

where z represents the state, $\mu > 0$ the viscosity coefficient, $\gamma > 0$ is a dissipation coefficient, $x \in [0, 1]$, and $t > 0$. We consider this problem in $D(A) = H^2_{\text{per}}(0, 1)$, the completion of C^∞-periodic functions in $H^2(0, 1)$. The initial conditions are unknown and we seek to estimate the solution by performing state measurements

$$y(t) = \left(\int_{\Omega_1} z(t,x)\mathrm{d}x, \ldots, \int_{\Omega_p} z(t,x)\,\mathrm{d}x\right)^T =: Cz(t) \tag{38}$$

in $\mathbb{R}^p$. To write (37) in the form of equation (1), we define

$$Az = \mu\frac{\partial^2 z}{\partial x^2} - \gamma z, \tag{39}$$

and

$$h(z, u) = -z\frac{\partial z}{\partial x}. \tag{40}$$

We consider $u \equiv 0$ for this nominal experiment, so we can also ignore the B operator. In the sections below, we show the problem of building a low-dimensional observer for the damped

Burgers equation (37) fits within our robust estimation framework. After some preliminary results describing the solutions to (37), we show that A generates a C_0-semigroup on $\mathcal{H}$ and that h satisfies the local Lipschitz-like condition of Assumption 3.2. This will be followed by numerical tests that demonstrate the performance of the nominal observer; the observer under the presence of a bounded uncertainty satisfying Assumption 4.1; and the auto-tuning implementation of the observer.

6.1 Theoretical justification

We first show that solutions to (37) are bounded in $\mathcal{H}$.

Lemma 6.1 (Solutions to (37) are bounded): *Let $z(t,\cdot)$ be a solution to the damped Burgers equation (37) with $z(0,\cdot) = z_0(\cdot) \in H^2_{\text{per}}(0,1)$. Then $\|z(t,\cdot)\|_{\mathcal{H}}$ remains bounded on any fixed time interval $(0,t_f)$.*

Proof: Multiplying equation (37) by $z(t,\cdot)$ and integrating over the periodic domain $(0,1)$ leads to

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_0^1 \frac{1}{2} z^2(t,x)\,\mathrm{d}x = &-\int_0^1 z^2(t,x)\frac{\partial z}{\partial x}(t,x)\,\mathrm{d}x\\ &-\mu\int_0^1\left(\frac{\partial z}{\partial x}(t,x)\right)^2\mathrm{d}x\\ &-\gamma\int_0^1 z^2(t,x)\,\mathrm{d}x.\end{aligned}$$

The first term on the right-hand-side above can be integrated and vanishes by periodicity. The second term results from integration-by-parts with the boundary terms vanishing by periodicity. Multiplying the remainder by an integrating factor leads to the following

$$\begin{aligned}\int_0^1 z^2(t,x)\,\mathrm{d}x = {}&\exp(-2\gamma t)\int_0^1 z_0^2(x)\,\mathrm{d}x\\ &-2\mu\int_0^t \exp(2\gamma(s-t))\int_0^1 z_x^2(s,x)\,\mathrm{d}x\,\mathrm{d}s.\end{aligned}$$

Since the last term above is always non-positive, we have shown that $\|z(t,\cdot)\|_{\mathcal{H}}$ decreases over time. ■

Theorem 6.2: *Let $\mathcal{D}(A) = H^2_{\text{per}}(0,1) \subset \mathcal{H}$ and $A : \mathcal{D}(A) \to \mathcal{H}$ be defined as in (39) with $\mu, \gamma > 0$. Then A generates a C_0-semigroup on $\mathcal{H}$.*

Proof: The operator A is dissipative as integration-by-parts leads to $\langle Az, z\rangle = -\langle z_x, z_x\rangle - \gamma\langle z, z\rangle \le 0$. Since A is densely defined, it generates a C_0-semigroup. ■

Corollary 6.3 (Stability of A_c): *If we compute the operator $A_c : \hat{\mathcal{H}} \to \hat{\mathcal{H}}$ using (19), then A_c generates an exponentially stable semigroup.*

Proof: If we consider $A_\mu z \equiv \mu z_{xx}$, the arguments made in Section 8.2 in Pazy (1983) for this periodic case show that A_μ is the infinitesimal generator of an analytic semigroup $T(t)$ satisfying $\|T(t)\| \le M$ for some $M \ge 1$ depending on the parameter μ. The semigroup $S(t) = \exp(-\gamma t)T(t)$ is generated by $A = A_\mu - \gamma z$, and is an analytic semigroup of solutions satisfying the bound $\|S(t)\| \le M\exp(-\gamma t)$.

Using (19), we have $A_c = \mathcal{T}^{\dagger *} A \mathcal{T}^{\dagger}$ and can show that for any $\hat{z} \in \hat{\mathcal{H}}_r$ with $\|\hat{z}\|_{\hat{\mathcal{H}}} = 1$, the following holds: $\langle A_c\hat{z}, \hat{z}\rangle = \langle A\mathcal{T}^{\dagger}\hat{z}, \mathcal{T}^{\dagger}\hat{z}\rangle \le \max_{\|z\|_{\mathcal{H}}=1}\langle Az, z\rangle$ since $\|\mathcal{T}^{\dagger}\hat{z}\| \le 1$. The operator A is self-adjoint, this implies the A_c generates a semigroup $S_c(t)$ satisfying the bound $\|S_c(t)\| \le M\exp(-\gamma t)$. ■

For functions that are piecewise differentiable, we can differentiate (37) with respect to x. By following the arguments of Lemma 1, multiplying the differentiated equation by $\frac{\partial z}{\partial x}$ instead, leads to the complex result that the spatial derivative, also known as the *enstrophy* $\|\frac{\partial z}{\partial x}(t,\cdot)\|_{\mathcal{H}}$, remains bounded on any fixed time interval $(0,t_f)$, cf. Pelinovsky (2012). Indeed, the additional $-\gamma u$ term limits the rate of growth over the usual estimates. The result is that $\|z(t,\cdot)\|_{H^1}$ remains bounded. This allows us to consider a local Lipschitz condition plus constant for (40) since

$$\begin{aligned}\|h(z_1) - h(z_2)\|_{\mathcal{H}} &\le \left(\|z_1\|_{H^1} + \|z_2\|_{H^1}\right)\|z_1 - z_2\|_{H^1}\\ &\le L_z(\|z_1 - z_2\|_{\mathcal{H}} + |z_1 - z_2|)\\ &\le L_z\|z_1 - z_2\|_{\mathcal{H}} + \Delta h,\end{aligned}$$

where $L_z = (\|z_1\|_{H^1} + \|z_2\|_{H^1})$ and $\Delta h = L_z|z_1 - z_2|$ where $|z_1 - z_2|$ is the H^1-seminorm, e.g. Brezis (1999, p. 121). We then use Lemma 1 together with the fact that the enstrophy is bounded (Pelinovsky, 2012).

6.2 Numerical tests

We consider here the case of the Burgers equation (37), with $\mu = 5\times10^{-3}$, $\gamma = 5\times10^{-2}$, boundary conditions $z(0,t) = z(1,t)$, and the initial condition:

$$z_0(x) = \begin{cases}0.5\sin(2\pi x), & x \in [0,.5],\\ 0, & x \in].5,1].\end{cases}$$

6.2.1 Nominal case

We first test the nominal case where there are no uncertainties explicitly added to the model (21), i.e. $\Delta h \equiv 0$. We report in Figure 1 the exact solution. We assume that we have access to 5 measurements centred at the following sensors locations: [0.15 0.35 0.55 0.75 0.95] with $|\Omega_i| = 0.1$, i.e. $\Omega_1 = [0.15-0.05,\ 0.15+0.05]$, $\Omega_2 = [0.35-0.05,\ 0.35+0.05]$, $\Omega_3 = [0.55-0.05,\ 0.55+0.05]$, $\Omega_4 = [0.75-0.05,\ 0.75+0.05]$, and $\Omega_5 = [0.95-0.05,\ 0.95+0.05]$. The corresponding measurements are plotted in Figure 2. We first implement the nominal observer (2), with the POD-based design (in Section 3.1). We use a POD basis of dimension 5, and discretise the PDE with linear finite elements resulting in an approximate state of dimension 64. Note that this number of sensors and discretisation dimension leads to the residual computation error $\|\text{res}_{Sylvester}\|_F = 0.0258$, which together with the maximum norm of z, $\max\|z\|_{\mathcal{H}} = 0.0625$, leads to the upper-bound $\|\text{res}_{Sylvester} z\| \le 0.0016$, this small error does not change the exponential convergence results of the observer, as discussed in Remark 3.6. We also introduce an initial condition error of 50%.

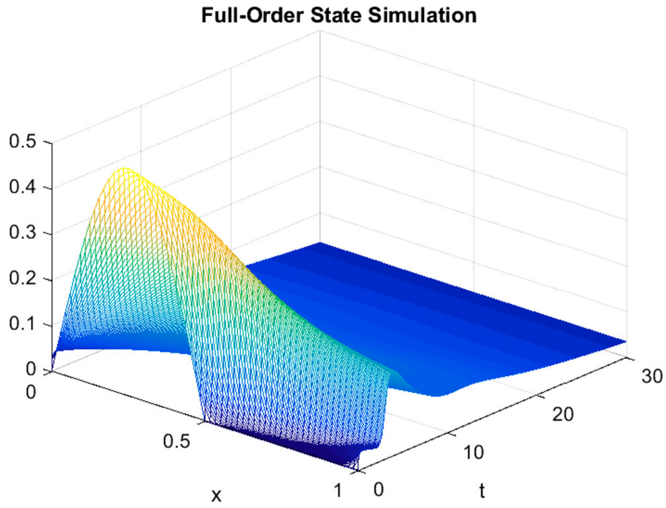

Figure 1. Exact state evolution.

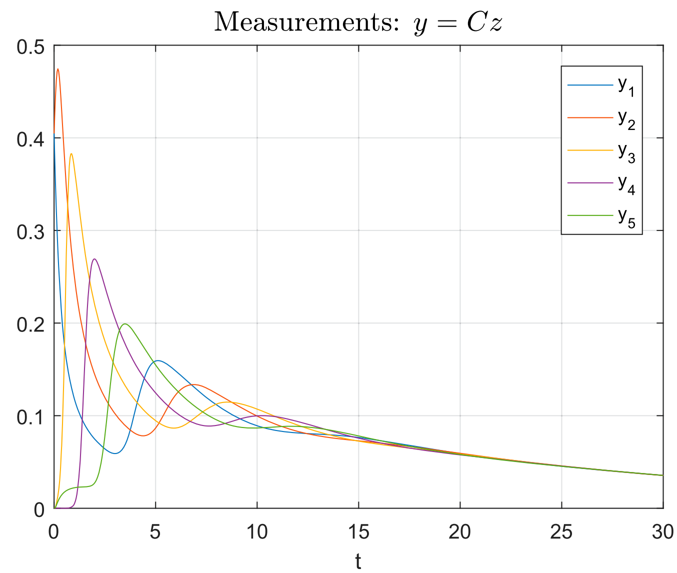

Figure 2. Output measurements: nominal case.

The estimate $\mathcal{T}^{\dagger}\hat{z}$ of the PDE solution z is shown in Figure 3. The estimated solution from only 5 measurements tracks toward the exact flow. The error between the estimate and the exact solution is reported in Figure 4, where we can see that the maximum error happens in the transient phase, due to the initial condition mismatch.

6.2.2 The case with parametric uncertainty

Next, to test the robustification term (24), we introduce an uncertainty in the viscosity coefficient $\delta\mu = -45 \times 10^{-4}$. We run again the nominal observer (2), without the robustification term. The corresponding estimated solution, and estimation error are given in Figures 5 and 6, respectively. We can see that the observer converges but the estimation error is larger than in the nominal case, due to the parametric uncertainty. Now, we test the robust observer (2), (7), and (24), where we select the gain to be $k = -10^3$. We see the clear effect of the robustification term in Figures 7 and 8. The estimation error rapidly decreases to zero, due to the robustification term that compensates for the model uncertainty.

6.2.3 An uncertain case with gain auto-tuning

We now present a test case with uncertainty in the viscosity coefficient. However, we do not 'settle' with our initial 'guess' of the observer gain k. Instead, we use the auto-tuning algorithm proposed in Section 5: implementing the auto-tuning ES algorithm presented in Theorem 5.1 with the learning cost function (34). We consider a simulation time $T = 30$ sec, to include the transient as well as the steady-state part of the estimation error. To motivate the need for auto-tuning, we first show the evolution of the learning cost function (34) as function of the observer again k. We report in Figure 9, the cost vs. gain plot, where we see that the constant value $k = -1 \times 10^3$ used in our first test, is not the optimal gain value. Indeed, the estimation performance, as defined by the learning cost, is optimal for a gain in the interval $[-300, -200]$. To ensure that the optimal gain for output error-based cost (34) is also optimal for the full state estimation error (i.e. Equation (34) where C is replaced with the identity matrix) we plot the full-state cost as function of the gain k in Figure 10. One clearly observes that the optimal gain for the output-based

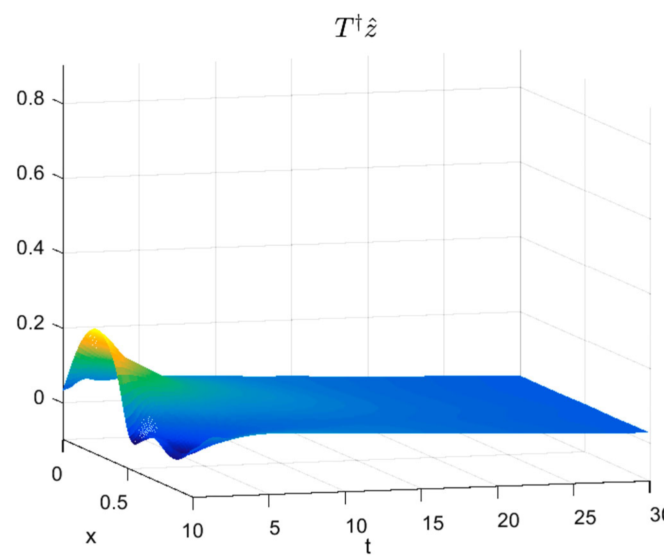

Figure 3. Estimated velocity: nominal case.

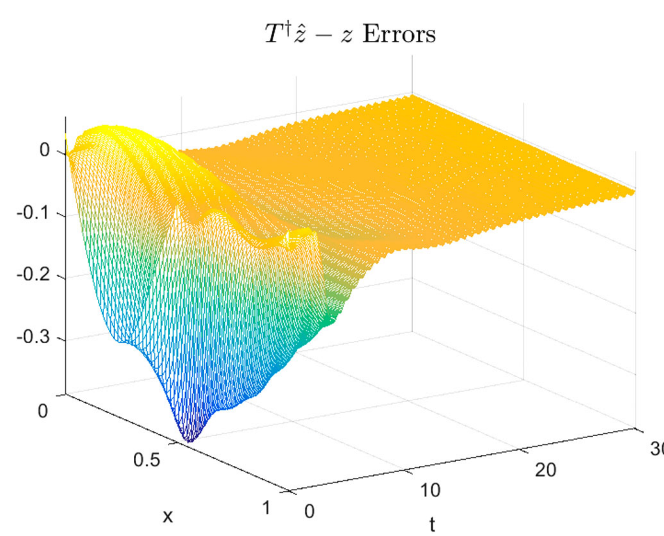

Figure 4. Estimation error: nominal case.

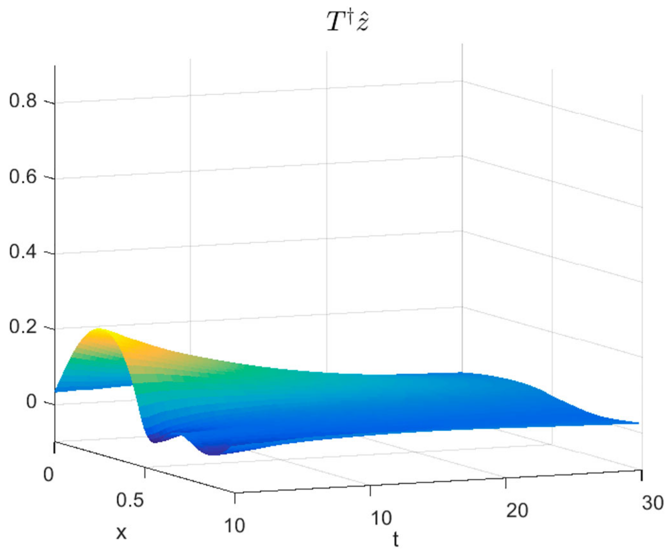

Figure 5. Estimated velocity: uncertain case with non-robust observer.

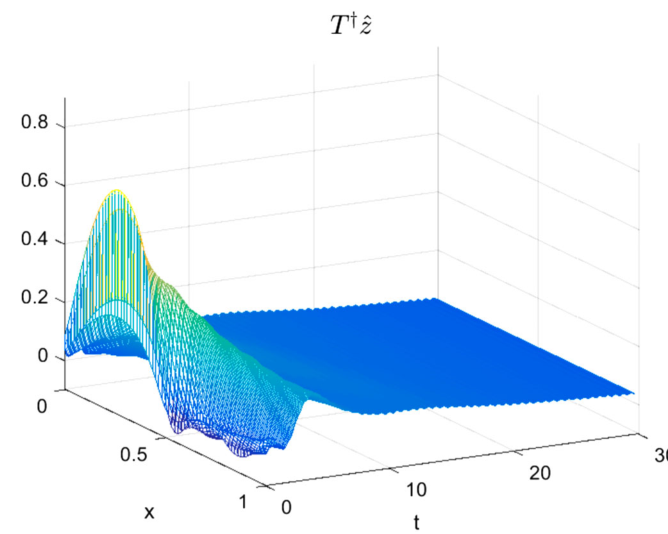

Figure 7. Estimated velocity: uncertain case with robust observer.

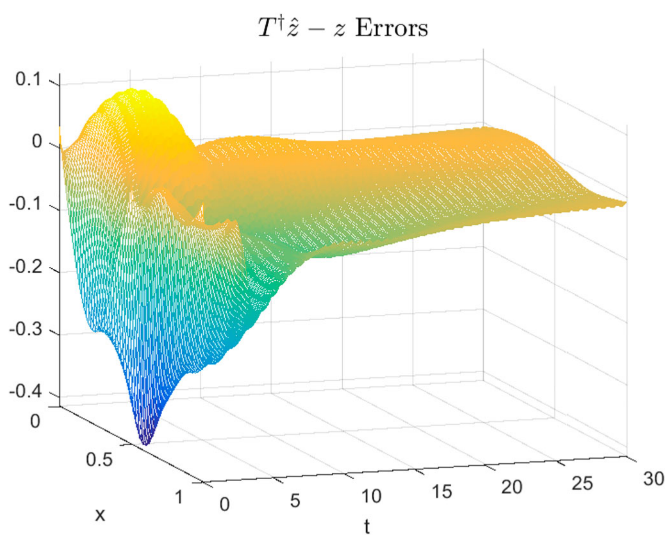

Figure 6. Estimation error: uncertain case with non-robust observer.

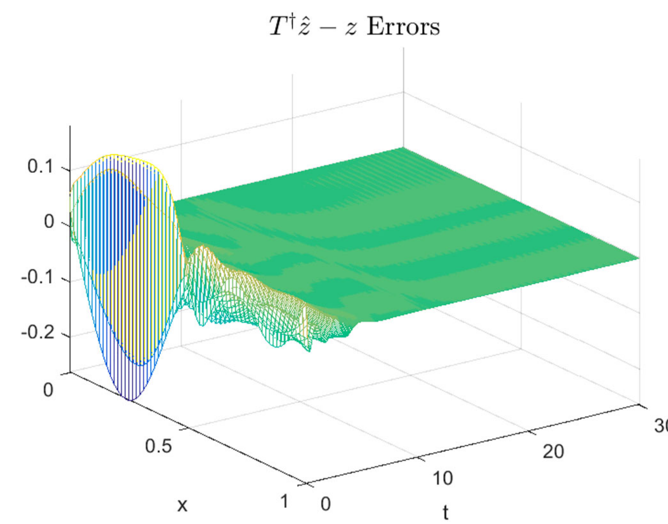

Figure 8. Estimation error: uncertain case with robust observer.

cost is also optimal for the full state-based cost. Thus, the gain k obtained through auto-tuning of the output-based learning cost will also have improvement in the overall estimation error for the entire state.

Next, we run the ES-based auto-tuning algorithm with the following constants: $a_k(0) = 10$, $\omega_k = 100$ [rad/sec], $\delta_k = 40$, and $\epsilon_k = 5 \times 10^{-4}$. The results of the auto-tuning are shown in Figures 11 and 12. We can see that the learning cost function decreases over the iterations and, as expected, the gains that provide the lowest estimation error are not necessarily the highest gains (in absolute value): the gain starts at -1×10^3 and converges to the neighbourhood of the optimal gain (within $[-300, -200]$).

We underline here that a classical extended Kalman filter approach has been applied to the same 1D Burgers problem in Borggaard et al. (2014). However, the extended Kalman filter does not handle parametric uncertainties. Furthermore, the Kalman filter would not be a good candidate for an

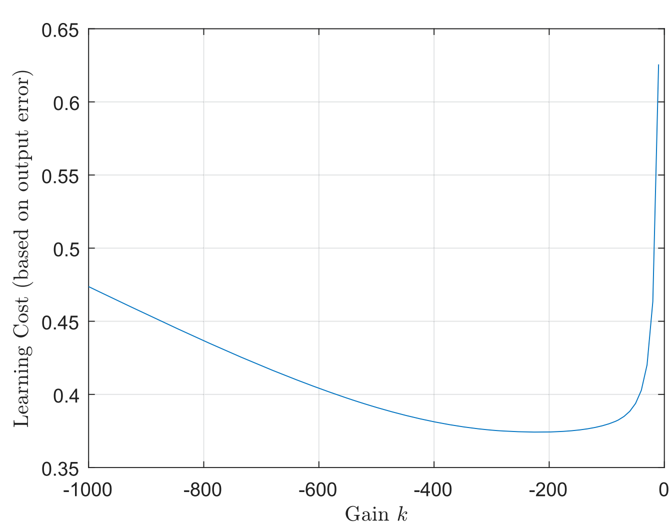

Figure 9. Learning (output-based) cost vs. gain.

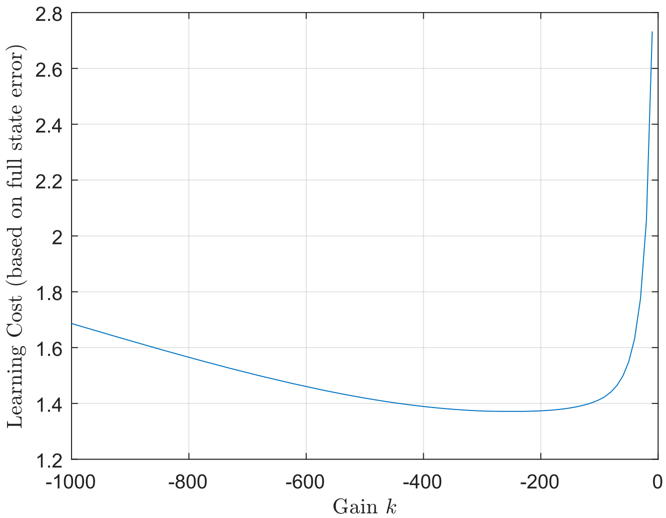

Figure 10. Learning (full state-based) cost vs. gain.

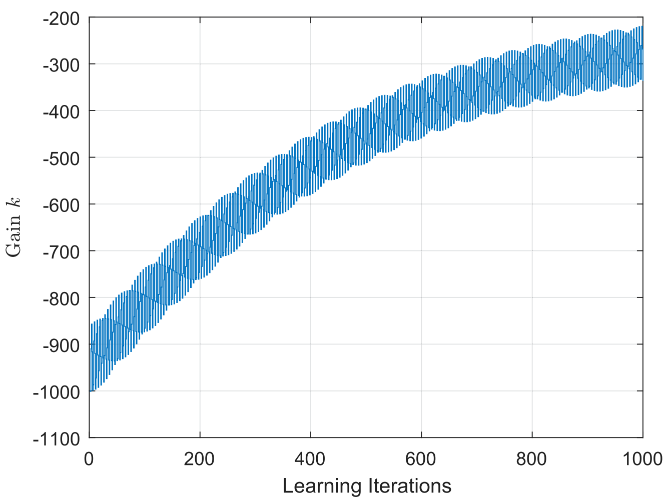

Figure 11. Gain vs. learning iterations: Uncertain case with robust observer.

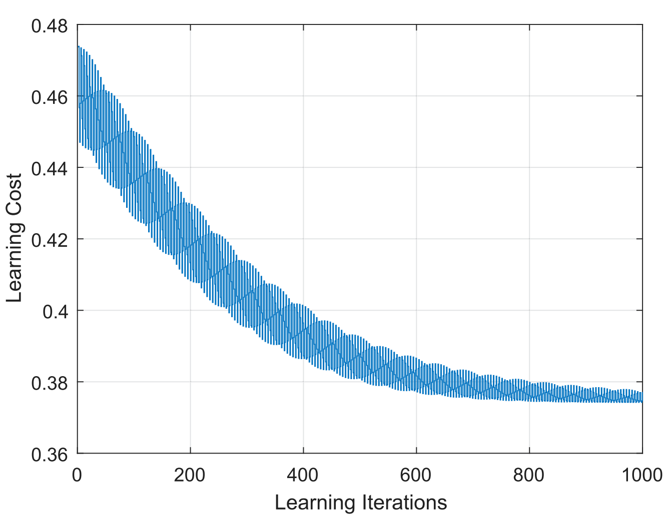

Figure 12. Learning cost vs. learning iterations: uncertain case with robust observer.

auto-tuning implementation, since its feedback gains have to satisfy algebraic Riccati equations, and cannot be easily learned online.

7. Conclusions

The problem of robust observer design for nonlinear infinite dimension systems is challenging. The results proposed in this paper are: (1) a robust reduced-order observer for nonlinear PDEs with bounded model uncertainties; (2) an IFT approach for online tuning of the observer gain; (3) an application to a non-trivial nonlinear PDE, namely the 1D Burgers equation.

For the large-scale discretisations required for complex nonlinear PDEs, it is infeasible to implement a full-order observer that can be reduced. Yet implementing an observer for a reduced-order model generally lacks theoretical justification. We have narrowed this gap in the current work by directly incorporating the model reduction subspaces within the observer design. Further studies will concern the case of model as well as measurement uncertainties. We intend to demonstrate the effectiveness of our approach on models where full-order observers are not feasible. For example, models that involve the 2D and 3D Boussinesq equations.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported in part by the National Science Foundation under grant DMS-1819110.

References

Atwell, J. A., Borggaard, J., & King, B. B. (2001). Reduced order controllers for Burgers' equation with a nonlinear observer. *International Journal of Applied Mathematics and Computer Science*, *11*, 1311–1330.

Balajewicz, M. J., Dowell, E. H., & Noack, B. R. (2013). Low-dimensional modelling of high-Reynolds-number shear flows incorporating constraints from the Navier-Stokes equation. *Journal of Fluid Mechanics*, *729*, 285–308.

Balas, M. J. (1981). Nonlinear state estimation and feedback control of nonlinear and bilinear distributed parameter systems. *Transactions of the ASME*, *102*, 78–82.

Battilotti, S. (2017). Robust observer design under measurement noise with gain adaptation and saturated estimates. *Automatica*, *81*, 75–86.

Benosman, M. (2016). Multi-parametric extremum seeking-based autotuning for robust input-output linearization control. *International Journal of Robust and Nonlinear Control*, *26*, 4035–4055.

Benosman, M., Borgaard, J., San, O., & Kramer, B. (2017). Learning-based robust stabilization for reduced-order models of 2D and 3D Boussinesq equations. *Applied Mathematical Modelling*, *49*, 162–181.

Benosman, M., & Lum, K. Y. (2010). Passive actuators' fault-tolerant control for affine nonlinear systems. *IEEE Transactions on Control Systems Technology*, *18*, 152–163.

Beutler, F. J. (1965). The operator theory of the pseudo-inverse. I. Bounded operators. *Journal of Mathematical Analysis and Applications*, *10*, 451–470.

Borggaard, J., Gugercin, S., & Zietsman, L. (2014). Compensators via H_2-based model reduction and proper orthogonal decomposition. *19th IFAC World congress* (pp. 7779–7784). South Africa, Cape Town.

Borggaard, J., Hay, A., & Pelletier, D. (2007). Interval-based reduced-order models for unsteady fluid flow. *International Journal of Numerical Analysis and Modeling*, *4*, 353–367.

Brezis, H. (1999). *Analyse fonctionnelle*. Paris: Dunod.

Burns, J. A., & Kang, S. (1990). *A control problem for Burgers equation with bounded input/output* (Technical report, ICASE, Report No. 90–45).

Feng, H., & Guo, B. Z. (2017). New unknown input observer and output feedback stabilization for uncertain heat equation. *Automatica*, *86*, 1–10.

Haddad, W. M., Chellaboind, V., & Nersesov, S. G. (2006). *Impulsive and hybrid dynamical systems: Stability, dissipativity, and control*. Princeton: Princeton University Press.

Hjalmarsson, H. (1998). Control of nonlinear systems using iterative feedback tuning. *IEEE American control conference* (pp. 2083–2087). Philadelphia, Pennsylvania.

Hjalmarsson, H. (2002). Iterative feedback tuning–an overview. *International Journal of Adaptive Control and Signal Processing*, *16*, 373–395.

Holmes, P., Lumley, J. L., & Berkooz, G. (1998). *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge: Cambridge University Press.

Jacobson, C. A., & Nett, C. N. (1988). Linear state-space systems in infinite-dimensional space: The role and characterization of joint stabilizability/detectability. *IEEE Transactions on Automatic Control*, *33*, 541–549.

Khalil, H. K. (1996). *Nonlinear systems* (2nd ed.). New York: Macmillan.

Kharkovskaya, T., Efimov, D., Polyakov, A., & Richard, J. P. (2018). Design of interval observers and controls for PDEs using finite-element approximations. *Automatica*, *93*, 302–310.

Killingsworth, N. J., & Krstic, M. (2006). PID tuning using extremum seeking. *IEEE Control Systems Magazine*, *26*, 70–79.

Koga, S., Benosman, M., & Borggaard, J. (2019). Learning-based robust observer design for coupled thermal and fluid systems. *IEEE American control conference*, to appear.

Krstic, M. (2000). Performance improvement and limitations in extremum seeking. *Systems & Control Letters*, *39*, 313–326.

Kunisch, K., & Volkwein, S. (2007). Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM Journal on Numerical Analysis*, *40*, 492–515.

Lequin, O., Gevers, M., Mossberg, M., Bosmans, E., & Triest, L. (2003). Iterative feedback tuning of PID parameters: Comparison with classical tuning rules. *Control Engineering Practice*, *11*, 1023–1033.

Miranda, R., Moreno, J. A., Chairez, J., & Fridman, L. (2012, January). Observer design for a class of hyperbolic PDE equation based on a distributed super twisting algorithm. *12th IEEE workshop on variable structure systems* (pp. 367–372). Mumbai, India.

Noack, B., Schlegel, M., Ahlborn, B., Mutschke, G., Morzynski, M., Comte, P., & Tadmor, G. (2008). A finite-time thermodynamics of unsteady fluid flows. *Journal of Non-Equilibrium Thermodynamics*, *33*, 103–148.

Pazy, A. (1983). *Semigroups of linear operators and applications to partial differential equations*, Vol. 44 of *Applied Mathematical Sciences*. New York, NY: Springer.

Pelinovsky, D. (2012). Sharp bounds on enstrophy growth in the viscous Burgers equation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *468*, 3636–3648.

Rim, D., & Mandli, K. T. (2018). *Model reduction of a parametrized scalar hyperbolic conservation law using displacement interpolation*. arXiv:1805.05938 [math.NA].

Schaum, A., Moreno, J. A., & Meurer, T. (2016). Dissipativity-based observer design for a class of coupled 1-D semi-linear parabolic PDE systems. *20th IFAC World congress* (pp. 98–103). Toulouse, France.

Smyshlyaev, A., & Krstic, M. (2010). *Adaptive control of parabolic PDEs*. Princeton: Princeton University Press.

Tan, Y., Nesic, D., & Mareels, I. (2006). On non-local stability properties of extremum seeking control. *Automatica*, *42*(6), 889–903.

Tan, Y., Nesic, D., Mareels, I., & Astolfi, A. (2009). On global extremum seeking in the presence of local extrema. *Automatica*, *45*(1), 245–251.

Wang, Z., Akhtar, I., Borggaard, J., & Iliescu, T. (2012). Proper orthogonal decomposition closure models for turbulent flows: A numerical comparison. *Computer Methods in Applied Mechanics and Engineering*, *237–240*, 10–26.

Witczak, M., Buciakowski, M., Puig, V., Rotondo, D., & Nejjari, F. (2016). An LMI approach to robust fault estimation for a class of nonlinear systems. *International Journal of Robust and Nonlinear Control*, *26*, 1530–1548.