2020 American Control Conference
Denver, CO, USA, July 1-3, 2020

Functional estimation of perturbed positive real infinite dimensional systems using adaptive compensators

Michael A. Demetriou

Abstract— **This paper extends earlier results on the adaptive estimation of nonlinear terms in finite dimensional systems utilizing a reproducing kernel Hilbert space to a class of positive real infinite dimensional systems. The simplest class of strictly positive real infinite dimensional systems has collocated input and output operators with the state operator being the generator of an exponentially stable C_0 semigroup on the state space X. The parametrization of the nonlinear term is considered in a reproducing kernel Hilbert space Q and together with the adaptive observer, results in an evolution system considered in $X \times Q$. Using Lyapunov-redesign methods, the adaptive laws for the parameter estimates are derived and the well-posedness of the resulting evolution error system is summarized. The adaptive estimate of the unknown nonlinearity is subsequently used to compensate for the nonlinearity. A special case of finite dimensional systems with an embedded reproducing kernel Hilbert space to handle the nonlinear term is also considered and the convergence results are summarized. A numerical example on a one-dimensional diffusion equation is considered.**

I. Introduction

The use of functional estimation of nonlinear terms in finite dimensional linear systems with online approximators was a well-studied topic in the late 80's and early 90's as can be observed in the inaugural and first issues of the IEEE Tr. on Neural Networks. Neural networks were used to approximate the nonlinear terms and the weights were trained (updated) using Lyapunov redesign methods.

More recently, [1] considered a class of MIMO nonlinear systems and formulated the adaptive parameter estimation of the assumed expansion of the nonlinear terms in a Euclidean space. This was subsequently extended to a functional space [2], [3], wherein the estimate of the unknown nonlinear function was embedded in a reproducing kernel Hilbert space (RKHS). The adaptive estimation scheme was then considered not in the cross product of Euclidean spaces but the cross product of the Euclidean space associated with the process dynamics and a Hilbert space associated with the functional estimate of the unknown nonlinear function.

The theory of RKHS was detailed in [4] and found its way to estimation and control of dynamical systems. In the work [5], [6], they utilized RKHS in order to adapt the centers of the radial basis functions (RBFs) used for the adaptive estimation of nonlinear terms.

The use of adaptive estimation techniques in infinite dimensional systems and in particular strictly positive real infinite dimensional systems has been addressed in [7] for systems with collocated input and output operators. The enabling assumption for using available signals, such as the process outputs and inputs, is that for the system under consideration to have a strictly positive real transfer function, [8]. In the time domain, this translates to the coupling of the solution to the Lyapunov equation to the input and output matrices (operators), [9]. It was subsequently extended to general strictly positive real infinite dimensional systems [10], [11], [12], [13]. An extension to the estimation of nonlinear terms was considered in [14] but it assumed that the nonlinear term admitted a series expansion.

This paper extends the earlier work on the use of RKHS for functional estimation of finite dimensional systems to a class of infinite dimensional systems. Two representative partial differential equations (PDEs) are used to illustrate the use of RKHS for functional estimation. The adaptive laws along with the conditions for parameter convergence are established and the well-posedenss of the resulting adaptive system are summarized. A numerical example of a diffusion PDE is provided in order to demonstrate the proposed adaptive functional estimation scheme used for the controller design of PDEs with nonlinear terms.

II. Motivating Examples: modeling and uncertainty parametrization

Two representative examples of a PDE in 1D are considered here, one with *in-domain* actuation and sensing

$$\begin{aligned} &\frac{\partial x(t,\xi)}{\partial t} = \frac{\partial^2 x(t,\xi)}{\partial \xi^2} + b(\xi;\xi_s)\Big(u(t)+w(t)+\gamma y(t)\Big), \\ &x(t,0)=x(t,\ell)=0,\;\; x(0,\xi)=x_0(\xi),\;\; 0\le\xi\le\ell, \\ &y(t)=\int_0^\ell b(\xi;\xi_s)x(t,\xi)\,\mathrm{d}\xi, \end{aligned} \tag{1}$$

and the other one with *boundary* actuation and sensing

$$\begin{aligned} &\frac{\partial x(t,\xi)}{\partial t} = \frac{\partial^2 x(t,\xi)}{\partial \xi^2},\;\; x(0,\xi)=x_0(\xi),\;\; 0\le\xi\le\ell, \\ &x(t,0)=0,\quad x_\xi(t,\ell)=u(t)+\gamma y(t)+w(t), \\ &y(t)=x(t,\ell). \end{aligned} \tag{2}$$

The spatially distributed state is denoted by $x(t,\xi)$, $\xi\in[0,\ell]$. The in-domain uncertainty term γ in (1) and the boundary uncertainty term γ in (2) may be an unknown constant that is desired to be identified. The source term w may be a constant term, a linear function of the output y, or, a nonlinear function of the output y. The first two choices for w can be incorporated into an affine function of the output y and this in turn can be absorbed by w being a nonlinear function of

M. A. Demetriou is with Aerospace Engineering, Worcester Polytechnic Institute, Worcester, MA 01609, USA, mdemetri@wpi.edu. The author gratefully acknowledges financial support from NSF-CMMI grant # 1825546.

the output y. Thus, (1) and (2) are re-written to reflect this

$$\begin{aligned}
&\frac{\partial x(t,\xi)}{\partial t} = \frac{\partial^2 x(t,\xi)}{\partial \xi^2} + b(\xi;\xi_s)\Big(u(t)+g(y(t))\Big),\\
&x(t,0) = x(t,\ell) = 0,\;\; x(0,\xi) = x_0(\xi),\;\; 0 \le \xi \le \ell, \qquad (1')\\
&y(t) = \int_0^\ell b(\xi;\xi_s)x(t,\xi)\,\mathrm{d}\xi,
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\partial x(t,\xi)}{\partial t} = \frac{\partial^2 x(t,\xi)}{\partial \xi^2},\;\; x(0,\xi) = x_0(\xi),\;\; 0 \le \xi \le \ell,\\
&x(t,0) = 0,\quad x_\xi(t,\ell) = u(t) + g(y(t)), \qquad (2')\\
&y(t) = x(t,\ell),
\end{aligned}$$

where

$$g(y) = a_1 + a_2 y + f(y), \tag{3}$$

with a_1, a_2 some unknown constants. The term $b(\xi;\xi_s)$ in (1) and ($1'$) represents the spatial distribution of the collocated actuator-sensor, parameterized by the spatial location ξ_s. In the event that this distribution is the Dirac delta function, then $b(\xi;\xi_s) = \delta(\xi - \xi_s)$. The system in (2) first considered in [7] in the context of adaptive estimation of collocated infinite dimensional systems was subsequently examined in detail in [15] in the context of non-square positive real infinite dimensional systems. It was presented as Example #1 in [15], and which had a symmetric version given by

$$\begin{aligned}
&\frac{\partial x(t,\xi)}{\partial t} = \frac{\partial^2 x(t,\xi)}{\partial \xi^2}, \qquad 0 \le \xi \le \ell,\\
&x_\xi(t,0) = u(t) - \gamma x(t,0) + w(t),\quad x(t,\ell) = 0,\\
&y(t) = x(t,0).
\end{aligned}$$

Both ($1'$) and ($2'$) can be placed in an abstract form, written as evolution equations in an appropriate Hilbert space. The definition of their state operator will differ and their input and output operators will have different representations; symbolically, they will be identical and both be members of a class of positive real infinite dimensional systems. This abstract representation for both ($1'$) and ($2'$) is

$$\begin{aligned}
\dot{x}(t) &= Ax(t) + B\left(u(t) + g(y(t))\right)\\
y(t) &= B^* x(t), \quad x(0) \in D(A),
\end{aligned} \tag{4}$$

where the state operator A will be defined over an appropriate Hilbert/Sobolev space, the input operator B provides information how the control signal $u(t)$ and the uncertainty/unmodelled dynamics $g(y(t))$ affect the state, and B^*, the adjoint of B, is the output operator and details the manner in which the sensing device obtains process measurements.

III. Adaptive estimation

Following the classification in [16], the state operator A is the infinitesimal generator of an exponentially stable C_0 semigroup $T_0(t), t \ge 0$ on the Hilbert space X, [17]. The rank 1 operator $B \in \mathcal{L}(\mathbb{R}, X)$ is the control input operator and the output operator is taken to be collocated to the input operator with $B^* \in \mathcal{L}(X, \mathbb{R})$. Following the notation in [16], the system in (4) is represented by the operator triple $\Sigma = (A, B, B^*)$.

The collocated system Σ is part of a special class of infinite dimensional systems termed *strictly positive real* (SPR). Following the definition of SPR systems in [14] with general output operators denoted by C, we can make the following assumption for $\Sigma = (A, B, C)$.

Assumption 1 (SPR system, [14]): Assume that the state operator A is the generator of an exponentially stable C_0 semigroup on X, $B \in \mathcal{L}(\mathbb{R}, X)$ and $C \in \mathcal{L}(X, \mathbb{R})$ such that there exist a nonnegative constant μ, a self-adjoint, nonnegative definite operator $P \in \mathcal{L}(X)$, and $Q \in \mathcal{L}(D(A), X)$ or $Q \in \mathcal{L}(D(A), \mathbb{R})$ such that for $\varphi \in D(A)$

$$(A + \mu I)^* P\varphi + P(A + \mu I)\varphi = -Q^* Q\varphi \tag{5a}$$

$$B^* P\varphi = C\varphi. \tag{5b}$$

The Lur'e equation in (5) is required for a more general class of SPR systems as it accounts for a general output operator not necessarily collocated with the input operator. An infinite dimensional system that satisfies (3) is then termed as an SPR (infinite dimensional) system. In the current class of systems considered, namely $\Sigma = (A, B, B^*)$, the collocated input output assumption with $C = B^*$ and A a dissipative operator satisfies (5a), (5b) with $P = I$. To formally state this, we make the following (relaxed) assumption for the class $\Sigma = (A, B, B^*)$, which satisfies the general assumption of SPR systems.

Assumption 2 (collocated SPR system): The system in (4), as represented by $\Sigma = (A, B, B^*)$, is assumed to have:

- The dissipative state operator A is the generator of an exponentially stable C_0 semigroup on X.
- The pair (A, B) is approximately controllable.

It should be noted that if (A, B) is approximately controllable, then (B^*, A^*) is approximately observable, [18]. The latter enables the design of an operator filter gain L such that $A - LC$ generates an exponentially stable C_0 semigroup.

For the remainder of the paper, Assumption 2 is considered to be satisfied. This assumption also applies to the case where the input operator is unbounded, as is the case in ($2'$), and in ($1'$) with $b(\xi;\xi_s) = \delta(\xi - \xi_s)$. In other words, a collocated input-output operator with the state operator being the generator of an exponentially stable C_0 semigroup satisfies the Lur'e equation (5a) which simplifies to $A + A^* < 0$. To accommodate for unbounded input and output operators, the spaces where they are defined must be enlarged. The Gelfand triple is considered $V \hookrightarrow X \hookrightarrow V^*$ with the embeddings dense and continuous. The state space X is the pivot space and V is a reflexive Banach space with V^* denoting its conjugate dual, [19]. To include the case where the process measurement is an m-dimensional vector, the output space is denoted by $\mathcal{Y} \in \mathbb{R}^m$. Since we assume a square system, then the input space $\mathcal{U}$ coincides with the output space. Then we have that the rank m input and output operators $B \in \mathcal{L}(\mathcal{Y}, V^*)$ and $B^* \in \mathcal{L}(V, \mathcal{Y})$, and the state operator $A \in \mathcal{L}(V, V^*)$. However, in view of the two representative examples ($1'$), ($2'$), we assume that the input and output operators are rank 1; i.e. $m = 1$. The control space $\mathcal{U}$ for square systems is identical to the output space $\mathcal{Y}$ with $\mathcal{U} = \mathcal{Y} = \mathbb{R}^1$.

Assumption 3: The class of systems represented by (4) is a square system where the input operator $B \in \mathcal{L}(\mathcal{Y}, V^*)$ is a rank 1 operator with $\mathcal{Y} \in \mathbb{R}^1$. Similarly, the output operator, taken to be the adjoint of the input operator, is a rank 1 output operator with $B^* \in \mathcal{L}(V,\mathcal{Y})$.

Finally, an assumption on the boundedness of the system (4) is required.

Assumption 4: The input signal u and the nonlinear function $g(y)$ are such that $\|x(t)\| \leq \gamma$, a.e. $t > 0$ for some $\gamma > 0$.

A. Parameter estimation

One way to estimate the unknown constant parameters a_1 and a_2 in (3), is to employ an adaptive observer. When $g(y)$ admits the particular series expansion

$$g(y(t)) = \sum_{i=1}^{N} a_i \phi_i(y(t)) \tag{6}$$

then one can set $\phi_1 = 1$ and $\phi_2(y) = y$ to arrive at a parametrization of $g(y)$ that absorbs (3). The functions ϕ_i can be polynomials of the output y or any other nonlinear functions of y but with the property that the semilinear dynamics $\dot{x} = Ax + Bg(B^*x)$ yield a well-posed system; for example a global Lipschitz continuity condition on $g(y)$ with $x_0 \in \mathcal{D}(A)$ can provide such a desired property, [20].

Following the parametrization (6), we can define the parameter space $\Theta \in \mathbb{R}^N$ associated with the unknown coefficients a_i, as the space of N-dimensional constant vectors with inner product and norm

$$\langle \theta, \chi \rangle_\Theta = \theta^T \chi, \quad \|\theta\|_\Theta^2 = \theta^T \theta, \quad \theta, \chi \in \Theta.$$

When the functional form of the uncertain term $g(y)$ is known, meaning that the nonlinear functions $\phi_i(\cdot)$, $i = 1,\ldots,N$ are known, then the adaptive observer for (4) with the parametrization (6) takes the form of the adaptive estimator in [14] and given by

$$\begin{aligned} \dot{\widehat{x}}(t) &= A\widehat{x}(t) + L\left(y(t) - C\widehat{x}(t)\right) + Bu(t) + B\sum_{i=1}^{N} \widehat{a}_i(t)\phi_i(y) \\ \dot{\widehat{a}}_i &= \delta_i \varepsilon(t)\phi_i(y(t)), \;\; \widehat{a}_i(0) = \widehat{a}_{i0}, \;\; i = 1,\ldots,N, \end{aligned} \tag{7}$$

where $\varepsilon(t) \triangleq y(t) - C\widehat{x}(t)$ is the output estimation error, $\widehat{a}_i(t)$ is the adaptive estimate of each a_i, $i = 1,\ldots,N$, $\widehat{x}(t)$ is the estimated state for (4), and $L \in \mathcal{L}(\mathcal{Y}, V^*)$ is the filter operator chosen so that $A - LC$ generates an exponentially stable C_0 semigroup on X. The constants $\delta_i > 0$, $i = 1,\ldots,N$ are termed the *adaptive gains*, [9]. It should be noted that the estimate of the unknown $g(y(t))$ is denoted by $\widehat{g}(t, y(t))$ and is given by the expansion (*cf.* (6))

$$\widehat{g}(t, y(t)) = \sum_{i=1}^{N} \widehat{a}_i(t)\phi_i(y(t)). \tag{8}$$

Setting the state error as $e(t) = x(t) - \widehat{x}(t)$, the parameter errors as $\widetilde{a}_i(t) = \widehat{a}_i(t) - a_i$ and bringing them in vector form

$$\widetilde{\theta}(t) = \widehat{\theta}(t) - \theta = \begin{bmatrix} \widehat{a}_1(t) - a_1 \\ \vdots \\ \widehat{a}_N(t) - a_N \end{bmatrix}, \; \Phi(y) = \begin{bmatrix} \phi_1(y) \\ \vdots \\ \phi_N(y) \end{bmatrix},$$

the state and parameter error dynamics resulting from (4) and (7) are considered in $X \times \mathbb{R}^N$ and are governed by

$$\begin{aligned} \dot{e}(t) &= (A - LC)e(t) + B\widetilde{\theta}^T(t)\Phi(y(t)) \\ \dot{\widetilde{\Theta}}(t) &= -\Delta\Phi(y(t))\varepsilon(t), \end{aligned} \tag{9}$$

where $\Delta > 0$ is the diagonal matrix of adaptive gains with $\{\Delta\}_{ii} = \delta_i$, $i = 1,\ldots,N$.

Similarly, one can define the *function estimate error*

$$\begin{aligned} \widetilde{g}(t, y(t)) &= g(y(t)) - \widehat{g}(t, y(t)) \\ &= \sum_{i=1}^{N} \left(a_i(t) - \widehat{a}_i(t)\right)\phi_i(y(t)) = \widetilde{\theta}^T(t)\Phi(y(t)). \end{aligned} \tag{10}$$

In terms of the representative PDE in $(1')$, the adaptive observer takes the specific form

$$\begin{aligned} \frac{\partial \widehat{x}(t,\xi)}{\partial t} &= \frac{\partial^2 \widehat{x}(t,\xi)}{\partial \xi^2} + b(\xi;\xi_s)u(t) + b(\xi;\xi_s)\sum_{i=0}^{N} \widehat{a}_i(t)\phi_i(y(t)) \\ &\quad + \mu(\xi)\left(y(t) - \int_0^\ell b(\xi;\xi_s)\widehat{x}(t,\xi)\,\mathrm{d}\xi\right) \\ \widehat{x}(t,0) &= \widehat{x}(t,\ell) = 0, \;\; \widehat{x}(0,\xi) = x_0(\xi), \;\; 0 \leq \xi \leq \ell, \\ \dot{\widehat{a}}_i(t) &= \delta_i\left(y(t) - \int_0^\ell b(\xi;\xi_s)e(t,\xi)\,\mathrm{d}\xi\right)\phi_i(y(t)) \end{aligned} \tag{11}$$

where $\mu(\xi)$ is the kernel representation of the adjoint of the filter operator L in (7). The error equations resulting from $(1')$ and (11) and associated with the formulation (9) are

$$\begin{aligned} \frac{\partial e(t,\xi)}{\partial t} &= \frac{\partial^2 e(t,\xi)}{\partial \xi^2} + b(\xi;\xi_s)\widetilde{\theta}^T(t)\Phi(y(t)) \\ &\quad - \mu(\xi)\int_0^\ell b(\xi;\xi_s)e(t,\xi)\,\mathrm{d}\xi \\ e(t,0) &= e(t,\ell) = 0, \;\; e(0,\xi) = e_0(\xi), \;\; 0 \leq \xi \leq \ell, \\ \dot{\widetilde{\theta}}(t) &= \left(\int_0^\ell b(\xi;\xi_s)e(t,\xi)\,\mathrm{d}\xi\right)\Delta\Phi(y(t)). \end{aligned} \tag{12}$$

The adaptive observer (7) corresponding to $(2')$ is given by

$$\begin{aligned} \frac{\partial \widehat{x}(t,\xi)}{\partial t} &= \frac{\partial^2 \widehat{x}(t,\xi)}{\partial \xi^2}, \quad \widehat{x}(0,\xi) = \widehat{x}_0(\xi), \quad 0 \leq \xi \leq \ell, \\ \widehat{x}(t,0) &= 0, \\ \widehat{x}_\xi(t,\ell) &= u(t) + \sum_{i=0}^{N} \widehat{\alpha}_i(t)\phi_i(y(t)) + \mu(\ell)\left(y(t) - \widehat{x}(t,\ell)\right), \\ \dot{\widehat{a}}_i(t) &= \delta_i\left(y(t) - \widehat{x}(t,\ell)\right)\phi_i(y(t)), \;\; \widehat{a}_i(0) = \widehat{a}_{i0}, \end{aligned} \tag{13}$$

and the associated error equations are given by

$$\begin{aligned} \frac{\partial e(t,\xi)}{\partial t} &= \frac{\partial^2 e(t,\xi)}{\partial \xi^2}, \quad e(0,\xi) = e_0(\xi), \quad 0 \leq \xi \leq \ell, \\ e(t,0) &= 0, \\ e_\xi(t,\ell) &= \widetilde{\theta}^T(t)\Phi(y(t)) - \mu(\ell)e(t,\ell), \\ \dot{\widetilde{\theta}}(t) &= e(t,\ell)\Delta\Phi(y(t)). \end{aligned} \tag{14}$$

The stability of the two error equations (12) and (14) can be established by equivalently examining their abstract representation (9). Using the Lyapunov candidate functional

$$V = |e(t)|_X^2 + \langle \widetilde{\theta}(t), \Delta^{-1}\widetilde{\theta}(t) \rangle_\Theta, \tag{15}$$

its derivative along (9) produces

$$\dot{V} = \langle e, (A - LC)e \rangle + \langle (A - LC)e, e \rangle \leq -\kappa |e|_X^2 \leq 0.$$

The convergence of the state error to zero can easily be established by applying the Bellman-Gronwall Lemma to (15). Most of the arguments follow the analysis in [14]. Similarly, the well-posedness follows the analysis presented in [21] when one applies the specific expression $B\Theta^T\Phi(y)$ for the nonlinear term. The error system (9) is examined in the space $X \times \Theta$ with the augmented state given by $z = (e, \widetilde{\theta})$ and the semilinear evolution operator $\mathcal{A}(t)$ on $X \times \Theta$ by

$$\mathcal{A}(t) = \begin{bmatrix} A_o & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & B[\cdot]^T\Phi(y) \\ -\Delta\Phi(y)\,(B^*[\cdot]) & 0 \end{bmatrix}, \quad (16)$$

where $A_o = A - LC$, resulting in (9) compactly written as the evolution system over the space $X \times \Theta$

$$\dot{z}(t) = \mathcal{A}(t)z(t), \quad z(0) = (e_0, \widetilde{\theta}(0)). \quad (17)$$

The remaining arguments use Galerkin approximation on a fixed point iterate along with the Lyapunov functional (15) to establish well-posedness, see [14] for details.

Remark 1: The matrix operator (16) has the skew adjoint structure that is characteristic of adaptive systems that first appeared in [22] for finite dimensional systems. Such a structure was first adapted to the infinite dimensional case [23], [21] and subsequently in the context of adaptive estimation of positive real infinite dimensional systems in [13]. The advantage of this skew adjoint structure is that it readily provides the condition for persistence of excitation. The term $-\Delta\Phi(y)\,(B^*[\cdot])$ in (17) can be replaced by

$$-\Delta\Phi(y)\,(B^*P[\cdot]) \longleftarrow -\Delta\Phi(y)\,(B^*[\cdot]),$$

when the general version of the positive real systems $\Sigma = (A,B,C)$ is considered as per Assumption 1. Following (5b), in this case the operator product B^*P can be replaced by the output operator C. This condition is of course the one that renders the adaptive laws (9) feasible since the update laws for $\widetilde{\Theta}$ utilize the available signals ε and $\Phi(y)$.

Similarly, one can express the condition for persistence of excitation needed to establish the parameter convergence $\lim_{t\to\infty}\widehat{\theta}(t) = \theta$. It requires the existence of T_0, δ_0 and ε_0 such that for each admissible $\psi \in \Theta$ with $|\psi|_\Theta = 1$ and sufficiently large $t > 0$, there exists $\widetilde{t} \in [t, t+T_0]$ such that

$$\left\| \int_{\widetilde{t}}^{\widetilde{t}+\delta_0} B\psi^T\Phi(y(\tau))\,\mathrm{d}\tau \right\|_{V^*} \geq \varepsilon_0. \quad (18)$$

Paving the way for the functional estimation and the subsequent norm-convergence of the functional learning presented in Section III-B below, one can consider a *weak* version of persistence of excitation as first presented in [24]. This entails the definition of the set

$$\overline{\Theta} = \Big\{\psi \in \Theta : \lim_{t\to\infty}\Big|\int_t^{t+\beta}\langle B\psi^T\Phi(y(\tau)), \varphi\rangle\,\mathrm{d}\tau\Big| = 0,\ \varphi \in V, \beta > 0\Big\} \quad (19)$$

and the ball $\mathcal{B}_\rho = \{\psi \in \Theta : |\psi|_\Theta \leq \sqrt{\rho}\}$. Then from (15) we have $\lim_{t\to\infty}|e(t)| = 0$ and

$$\lim_{t\to\infty}\operatorname{dist}(\widetilde{\theta}(t), \overline{\Theta}\cap\mathcal{B}_\rho) = 0. \quad (20)$$

Additionally, one has

$$\text{weak} - \lim_{t\to\infty}\widehat{\theta}(t) = \theta + \Pi_{\overline{\Theta}}(\widehat{\theta}(0) - \theta), \quad (21)$$

where $\Pi_{\overline{\Theta}}$ denotes the orthogonal projection of Θ onto $\overline{\Theta}$.

B. Functional estimation

Now, the nonlinear term $g(y)$ in (4) is no longer assumed to admit the expansion (6). We denote by Q the Hilbert space of functions defined on $\mathcal{Y}$, i.e. $f : \mathcal{Y} \to \mathbb{R}^1$, with the evaluation functional over Q which evaluates each function at a point $y \in \mathcal{Y}$ by

$$\lambda_y : f \to f(y), \quad \forall f \in Q, \quad (22)$$

i.e. $f(y) = \lambda_y(f)$. The argument of $f(y)$ can be "viewed" as the spatial variable as considered in the adaptive estimation of spatially varying parameters in PDEs. Through the appropriate construction of the kernels, defined below, one has that the evaluation functional λ_y is bounded and thus the Hilbert (parameter) space Q is a RKHS. Using the Riesz representation theorem [25] we have that for all $y \in \mathcal{Y}$, there is an element $\kappa(\cdot,\cdot) : \mathcal{Y}\times\mathcal{Y} \to \mathbb{R}^1$ with $\kappa_y = \kappa(y,\cdot)$ (reproducing kernel) that has the reproducing property

$$f(y) = \lambda_y(f) = \langle f, \kappa(y,\cdot)\rangle_Q = \langle f, \kappa_y\rangle_Q, \forall f \in Q\,\forall y \in \mathcal{Y}, \quad (23)$$

Following [26], the inner product enables the evaluation of the kernel function at "spatial" points of the *data space* $\mathcal{Y}$. For "spatial" points $y_i, y_j \in \mathcal{Y}$ and corresponding elements in the *feature space* $f(y_i), f(y_j) \in Q$, then

$$\langle f(y_i), f(y_j)\rangle_Q = \langle \kappa(y_i,\cdot), \kappa(y_j,\cdot)\rangle_Q = \kappa(y_i, y_j).$$

This is known as the *kernel trick* [26], and reduces the calculations in high-dimensional spaces.

Essential to the derivation of the update laws via Lyapunov-redesign method, is the definition of the *adjoint of the evaluation functional* $\lambda_y^* : \mathcal{Y} \to Q$ given by

$$\langle \varepsilon, \lambda_y(f)\rangle_{\mathcal{Y}} = \langle \varepsilon\kappa_y, f\rangle_Q = \langle \lambda_y^*(\varepsilon), f\rangle_Q. \quad (24)$$

Using (23), we detail the steps used for the extraction of the update laws. Consider the term $B\lambda_y(f)$ in weak form

$$\begin{aligned} \langle \varphi, B\lambda_y(f)\rangle_{V,V^*} &= \langle B^*\varphi, \lambda_y(f)\rangle_{\mathcal{Y}} \\ &= \langle \lambda_y^*(B^*\varphi), f\rangle_Q \end{aligned}, \quad \varphi \in V, f \in Q. \quad (25)$$

Consider now the abstract representation (4) where we no longer assume the expansion (6)

$$\dot{x}(t) = Ax(t) + Bu(t) + B\lambda_{y(t)}(g), \quad \text{in } V^*. \quad (26)$$

The expression for the adaptive observer (7) is now given by

$$\dot{\widehat{x}}(t) = A\widehat{x}(t) + L(y(t) - C\widehat{x}(t)) + Bu(t) + B\lambda_{y(t)}(\widehat{g}), \quad (27)$$

where $\widehat{g}$ denotes the time varying (adaptive) estimate of g. Using (26) and (27), the evolution equation of the state estimation error is given by

$$\dot{e}(t) = (A - LC)e(t) + B\lambda_{y(t)}(\widetilde{g}). \quad (28)$$

In order to extract the adaptation for $\widehat{g}$, we consider the Lyapunov functional (*cf.* (15))

$$V(e, \widetilde{f}) = |e(t)|_X^2 + \langle \mathcal{G}^{-1}\widetilde{g}, \widetilde{g}\rangle_Q \quad (29)$$

where $\mathcal{G} \in L(Q,Q)$ is the positive definite, self-adjoint linear operator [1], taking the role of the adaptive gain matrix Δ in (15). Taking the derivative along the error equation (28) then

$$\dot{V} = \langle e, A_o e\rangle + \langle A_o e, e\rangle + 2\langle e, B\lambda_{y(t)}(\widetilde{g})\rangle + 2\langle \mathcal{G}^{-1}\dot{\widetilde{g}}, \widetilde{g}\rangle_Q. \quad (30)$$

Using (25), the term $2\langle e, B\lambda_{y(t)}(\widetilde{g})\rangle$ is equivalent to

$$2\langle e, B\lambda_{y(t)}(\widetilde{g})\rangle_{V,V^*} = 2\langle \lambda_{y(t)}^*(\varepsilon), \widetilde{g}\rangle_Q,$$

where $\varepsilon(t) = B^*e(t)$ is the output estimation error. Using the

[1]One can ignore $\mathcal{G}$ since $\langle p, \mathcal{G}q\rangle_Q$, $p,q \in Q$ can be viewed as a redefinition of the inner product in Q and thus $\langle p, \mathcal{G}q\rangle_Q \simeq \langle p,q\rangle_Q$

above, the derivative of V simplifies to

$$\dot{V} = \langle e, A_o e\rangle + \langle A_o e, e\rangle + 2\langle \lambda^*_{y(t)}(\varepsilon), \widetilde{g}\rangle_Q + 2\langle \mathcal{G}^{-1}\dot{\widetilde{g}}, \widetilde{g}\rangle_Q.$$

The adaptive laws for $\widetilde{g}$ can now be extracted and given by

$$\dot{\widetilde{g}} = -\mathcal{G}\lambda^*_{y(t)}(\varepsilon(t)) \tag{31}$$

or in weak form

$$\langle \dot{\widetilde{g}}, p\rangle_Q = -\langle \mathcal{G}\lambda^*_{y(t)}(\varepsilon(t)), p\rangle_Q = -\langle \varepsilon(t), \lambda_{y(t)}(\mathcal{G}p)\rangle_Q \tag{32}$$

for all test functions $p \in Q$.

The counterpart of the abstract operator $\mathcal{A}$ in (16), defined now over $X \times Q$ is given by

$$\mathcal{A}(t) = \begin{bmatrix} A_o & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & B\lambda_{y(t)}([\cdot]) \\ -\mathcal{G}\lambda^*_{y(t)}(B^*[\cdot]) & 0 \end{bmatrix} \tag{33}$$

The extension of the adaptive functional estimation using RKHS's can be summarized in the following lemma.

Lemma 1: Assume that the PDEs under consideration can be represented in the form (4) where Assumptions 2, 3 and 4 are satisfied. Then the adaptive observer of the system in the form (26) is given by (27) with the adaptive rules for g given in weak form by (32). Additionally, the resulting error system, written as an evolution system in $X \times Q$

$$\frac{d}{dt}\begin{bmatrix} e(t) \\ \widetilde{g}(t) \end{bmatrix} = \mathcal{A}(t)\begin{bmatrix} e(t) \\ \widetilde{g}(t) \end{bmatrix} \tag{34}$$

with $\mathcal{A}(t)$ given by (33), is well-posed and $\lim_{t\to\infty} |e(t)|_X = 0$. Furthermore, if there exist T_0, δ_0 and ε_0 such that for each admissible element $q \in Q$ with $|q|_Q = 1$ and sufficiently large $t > 0$, there exists $\widetilde{t} \in [t, t+T_0]$ such that

$$\left\| \int_{\widetilde{t}}^{\widetilde{t}+\delta_0} B\lambda_{y(\tau)}(q)\, \mathrm{d}\tau \right\|_{V^*} \geq \varepsilon_0. \tag{35}$$

then $\lim_{t\to\infty} \|\widehat{g}(t) - g\|_Q = 0$.

Proof: Assumption 4 ensures that (26) coincides with the definition of a *plant* in [23]. Using (34) where the time-dependent evolution operator $\mathcal{A}(t)$ is given by (33) one can immediately use Theorem 3.2 and Theorem 3.4 in [23] to establish well-posedness and convergence. ■

One can use the adaptive estimate of the nonlinear term to improve controller performance. Denote by $u_0(t)$ the nominal control signal that in the absence of the nonlinear term $g(y)$, provides the requisite performance; such a controller can be based on full state feedback given by $u_0(t) = -Kx(t)$ where $K \in \mathcal{L}(V, \mathcal{U})$ is the full state feedback operator design to provide a certain performance. In the absence of full state measurements, one must use the state estimate to implement such a nominal controller. The state of the adaptive observer can serve as such and the adaptive controller that includes the nonlinear compensation is given by

$$u(t) = -\widehat{g}(y) - K\widehat{x}(t) \tag{36}$$

The resulting closed-loop system is given by

$$\dot{x}(t) = (A - BK)x(t) + BKe(t) + B\widetilde{g}(y). \tag{37}$$

When examining the stability and well-posedness of (37), all three equations must be considered in the space $X \times X \times Q$ with aggregate state $\zeta(t) = (e(t), x(t), \widetilde{g}(t)))$

$$\dot{\zeta} = \begin{bmatrix} A_o & 0 & B\lambda_{y(t)}([\cdot]) \\ BK & A - KC & B\lambda_{y(t)}([\cdot]) \\ -\mathcal{G}\lambda^*_{y(t)}(B^*[\cdot]) & 0 & 0 \end{bmatrix}\zeta \tag{38}$$

Lemma 2 (CL system): Assume that the square infinite dimensional system (26) satisfies Assumptions 2–4. Then the adaptive compensator given by (27), (32) and (36) results in a well-posed system described by (38) with

$$\lim_{t\to\infty} |x(t)|_X = 0 \text{ and } \lim_{t\to\infty} |e(t)|_X = 0.$$

Proof: The aggregate system (38) is in the form of Theorem 2.9 in [21] and therefore the well-posedness of (38) can be established using implicit function theorem. Similarly, Theorems 3.4 and 3.6 in [21] provide the convergence. ■

C. Special case: process dynamics in $\mathbb{R}^n$

When the process dynamics (4) is now governed by a finite dimensional system in $\mathbb{R}^n$ with the Gelfand triple collapsing and thus $V = X = V^* = \mathbb{R}^n$, the conditions for persistence of excitation are simplified and the convergence of the state estimation error can be easily established. Thus the finite dimensional counterpart of (26) is now given by

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + B\lambda_{y(t)}(g), \quad \text{in } \mathbb{R}^n, \\ y(t) &= Cx(t) \end{aligned} \tag{39}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$ and $C \in \mathbb{R}^{1\times n}$. We can impose the general condition for SPR systems, thus requiring the triple $\Sigma = (A, B, C)$ satisfy the Lur'e matrix equations (*cf.* (5))

$$A^T P + PA = -Q^T Q, \qquad B^T P = C. \tag{40}$$

The associated adaptive observer is symbolically identical to the infinite dimensional counterpart in (27) and given by

$$\dot{\widehat{x}}(t) = A\widehat{x}(t) + L(y(t) - C\widehat{x}(t)) + Bu(t) + B\lambda_{y(t)}(\widehat{g}), \tag{41}$$

where $\widehat{x}(t) \in \mathbb{R}^n$ is the adaptive estimate of the state vector $x(t)$ and $\widehat{g}$ is still the adaptive estimate of g. The associated error equation is given by

$$\dot{e}(t) = (A - LC)e(t) + B\lambda_{y(t)}(\widetilde{g}). \tag{42}$$

The Lyapunov functional needed to extract the adaptive laws is given by

$$V(e, \widetilde{f}) = e^T(t)Pe(t) + \langle \mathcal{G}^{-1}\widetilde{g}(t), \widetilde{g}(t)\rangle_Q \tag{43}$$

with $\mathcal{G} \in \mathcal{L}(Q, Q)$ the adaptive gain operator. Its derivative along the trajectories of the state estimation error is

$$\begin{aligned} \dot{V} &= e^T P A_o e + e^T A_o^T P e + 2e^T P B\lambda_{y(t)}(\widetilde{g} + 2\langle \mathcal{G}^{-1}\dot{\widetilde{g}}, \widetilde{g}\rangle_Q \\ &= -e^T Q^T Q e + 2\varepsilon^T \lambda_{y(t)}(\widetilde{g}) + 2\langle \mathcal{G}^{-1}\dot{\widetilde{g}}, \widetilde{g}\rangle_Q \end{aligned} \tag{44}$$

Using (24), (25), the third term is

$$\begin{aligned} \varepsilon^T \lambda_{y(t)}(\widetilde{g}) &= \langle \varepsilon, \lambda_{y(t)}(\widetilde{g})\rangle_{\mathcal{Y}} \\ &= \langle \varepsilon \kappa_{y(t)}, \widetilde{g}\rangle_Q = \langle \lambda^*_{y(t)}(\varepsilon), \widetilde{g}\rangle_Q \end{aligned}$$

which immediately yields the same adaptive laws as in (32).

Similar to the infinite dimensional case, the adaptive controller can be used to minimize/reject the effects of the unknown nonlinear term. The controller has the same structure as the infinite dimensional one (35).

Lemma 3: Assume that the triple $\Sigma = (A, B, C)$ in (39) satisfies the Lur'e equations (40). Then the compensator (41),

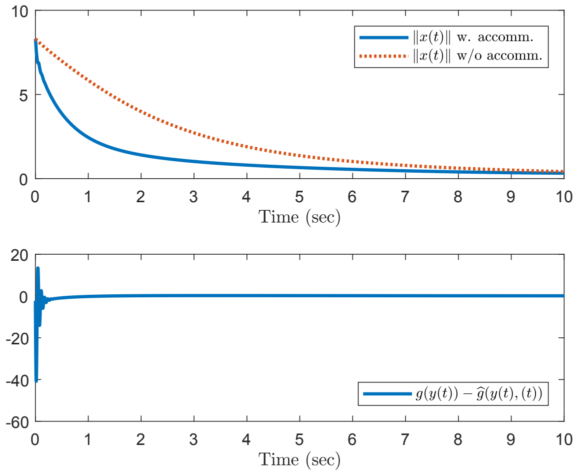

Fig. 1. Evolution of state and functional norms.

(32) and (36) results in a well-posed system in $\mathbb{R}^N \times \mathbb{R}^N \times Q$ with boundedness of all signals and asymptotic convergence of the plant state $x(t)$ and estimation error $e(t)$ to zero. The proof is based on the proofs of Lemmas 1, 2 adjusted for $V = X = V^* = \mathbb{R}^N$.

IV. Numerical Example and Conclusions

The PDE in (1′), modified to include a thermal diffusivity parameter, is considered in $[0,\ell] = [0,1]$

$$x_t(t,\xi) = a x_{\xi\xi}(t,\xi) + \delta(\xi - \xi_s)\left(u(t) + g(y(t))\right)$$

with $a = 0.01$ and the actuator location $\xi_s = 0.215\ell$. The nonlinear term was set as $g(y) = 2 \times 10^{-4} y^3$. The controller (36) used an operator gain K based on LQR design with penalties given by $\|x(t)\|^2$ and $100u^2(t)$ and the adaptive estimate of $g(y)$. The filter used $\mu(\xi) = 0.001\delta(\xi - \xi_s)$. Radial basis functions (RBFs) were selected for the functional estimation, with $\kappa_y(q) = \exp\{-\frac{|y-q|^2}{2\sigma^2}\}$. The standard deviation was selected as $\sigma = \frac{100}{2\sqrt{\log(2)}}$ and the means were evenly distributed in the interval $[-10, 20]$. For the approximation of $g(y)$, using finite dimensional subspaces $Q^N \subset Q$, a total number $N = 31$ of RBFs were used $g(y) \approx \sum_{i=1}^{N} \theta_i \kappa_{y_i}(\cdot)$ and for the approximation of (1′) a Galerkin-based scheme was used to approximate with a total of 50 linear elements.

Figure 1a depicts the evolution of the state norm with the accommodating controller (36) and with only using the nominal controller $u_0(t) = -K\widehat{x}(t)$. When the functional estimation is used in (36) the controller performance is significantly improved. Figure 1b depicts the evolution of the functional error $g(y(t)) - \widehat{g}(y(t),t)$ where it is observed that it also converges to zero.

The earlier work on the use of RKHS for functional estimation of finite dimensional systems has been extended to a class of infinite dimensional systems. The well-posedness and convergence of the adaptive observer was summarized and its use as a compensator was demonstrated in a 1D PDE.

References

[1] V. Stepanyan and A. Kurdila, "Asymptotic tracking of uncertain systems with continuous control using adaptive bounding," *IEEE Trans. on Neural Networks*, vol. 20(8), pp. 1320–1329, Aug 2009.

[2] P. Bobade, S. Majumdar, S. Pereira, A. J. Kurdila, and J. B. Ferris, "Adaptive estimation in reproducing kernel hilbert spaces," in *Proc. of the American Control Conf.*, May 2017, pp. 5678–5683.

[3] P. Bobade, S. Majumdar, S. Pereira, A. J. Kurdila, and J. B. Ferris, "Adaptive estimation for nonlinear systems using reproducing kernel Hilbert spaces," *Adv. Comput. Math.*, vol. 45(2), pp. 869–896, 2019.

[4] N. Aronszajn, "Theory of reproducing kernels," *Trans. Amer. Math. Soc.*, vol. 68, pp. 337–404, 1950.

[5] H. A. Kingravi, G. Chowdhary, P. A. Vela, and E. N. Johnson, "A reproducing kernel hilbert space approach for the online update of radial bases in neuro-adaptive control," in *Proc. of the IEEE Conf. on Decision and Control and European Control Conf.*, Dec 2011.

[6] ——, "Reproducing kernel hilbert space approach for the online update of radial bases in neuro-adaptive control," *IEEE Trans. on Neural Networks and Learning Systems*, vol. 23(7), July 2012.

[7] M. A. Demetriou and K. Ito, "Adaptive observers for a class of infinite dimensional systems," in *Proc. of the 13th World Congress, IFAC*, San Francisco, CA, July 1996.

[8] B. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis: A Modern Systems Theory Approach*, ser. Prentice Hall Network Series. Englewood Cliffs, NJ: Prentice Hall, 1973.

[9] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[10] R. Curtain, M. Demetriou, and K. Ito, "Adaptive observers for structurally perturbed infinite dimensional systems," in *Proc. of the IEEE Conference on Decision and Control*, Hyatt Regency, San Diego, CA, December 10-12 1997, pp. 509–514.

[11] R. F. Curtain, M. A. Demetriou, and K. Ito, "Adaptive observers for slowly time varying infinite dimensional systems," in *Proc. of the IEEE Conference on Decision and Control*, vol. 4, Dec 1998, pp. 4022–4027.

[12] M. A. Demetriou, R. F. Curtain, and K. Ito, "Adaptive observers for structurally perturbed positive real delay systems," in *Proc. of The Fourth Int'l Conference on Optimization: Techniques and Applications*, Curtin University of Technology, Perth, Australia, July 1-3 1998.

[13] R. F. Curtain, M. A. Demetriou, and K. Ito, "Adaptive compensators for perturbed positive real infinite-dimensional systems," *Int. J. Appl. Math. Comput. Sci.*, vol. 13(4), pp. 441–452, 2003.

[14] M. A. Demetriou, K. Ito, and R. C. Smith, "Adaptive techniques for the MRAC, adaptive parameter identification, and on-line fault monitoring and accommodation for a class of positive real infinite dimensional systems," *Internat. J. Adapt. Control Signal Process.*, vol. 23(2), pp. 193–215, 2009.

[15] M. A. Demetriou, "Adaptive observers for non-square positive real infinite dimensional systems," in *Proc. of the American Control Conf.*, Boston, MA, July 6-8 2016.

[16] ——, "Classification of adaptive observers for non-square infinite dimensional systems with structured perturbations and disturbances," in *Proc. of the American Control Conf.*, May 24-26 2017, pp. 493–498.

[17] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*. New York: Springer-Verlag, 1983.

[18] R. F. Curtain and H. J. Zwart, *An Introduction to Infinite Dimensional Linear Systems Theory*, ser. Texts in Applied Mathematics, Vol. 21. Berlin: Springer-Verlag, 1995.

[19] R. E. Showalter, *Hilbert Space Methods for Partial Differential Equations*. London: Pitman, 1977.

[20] V. Barbu, *Analysis and control of nonlinear infinite-dimensional systems*, ser. Mathematics in Science and Engineering. Academic Press, Inc., Boston, MA, 1993, vol. 190.

[21] M. Böhm, M. A. Demetriou, S. Reich, and I. G. Rosen, "Model reference adaptive control of distributed parameter systems," *SIAM J. Control Optim.*, vol. 36, no. 1, pp. 33–81, 1998.

[22] A. P. Morgan and K. S. Narendra, "On the stability of nonautonomous differential equations $\dot{x} = [A + B(t)]x$, with skew symmetric matrix $B(t)$," *SIAM J. Control Optimization*, vol. 15(1), pp. 163–176, 1977.

[23] J. Baumeister, W. Scondo, M. A. Demetriou, and I. G. Rosen, "On-line parameter estimation for infinite-dimensional dynamical systems," *SIAM J. Control Optim.*, vol. 35, no. 2, pp. 678–713, 1997.

[24] M. A. Demetriou and I. G. Rosen, "On the persistence of excitation in the adaptive estimation of distributed parameter systems," *IEEE Trans. on Automatic Control*, vol. 39, no. 5, pp. 1117–1123, May 1994.

[25] M. Reed and B. Simon, *Methods of modern mathematical physics. I. Functional analysis*. Academic Press, New York-London, 1972.

[26] K. Slavakis, S. Theodoridis, and I. Yamada, "Adaptive constrained learning in reproducing kernel hilbert spaces: The robust beamforming case," *IEEE Trans. Signal Process.*, vol. 57(12), Dec 2009.