Deadbeat Simultaneous Parameter-State Estimation for Linear Continuous-time Systems: a Kernel-based Approach

Peng Li, Francesca Boem, Gilberto Pin, Thomas Parisini

Abstract— The simultaneous estimation of parameters and states in continuous-time linear MISO systems is considered in this paper. The estimates of system's parameters and states are provided simultaneously by a single estimator, where Volterra operators are suitably applied to the I/O measurements. This makes the estimator independent on the initial conditions thanks to the design of suitably shaped kernel functions equipped with non-asymptotic properties. As a result, in the noise-free scenario, instantaneous convergence can be obtained. No high-gain injection nor periodic resetting is necessary. Numerical examples are reported showing the effectiveness of the proposed estimator.

I. INTRODUCTION

The identification of the parameters and states of dynamic systems is one of the fundamental problems in control engineering. The problem was first formulated as a nonlinear state observation problem by augmenting the parameters into the state vector and designing an observer for the nonlinear dynamics [1]. Among others, [2] simplifies the problem by using the extended Kalman filter and proposing a computationally economical joint parameter and state estimator. In [3], [4], parameter-state estimation is realized based on adaptive methods which are reviewed in details in [5]. Nevertheless, most of the classical methods can only provide asymptotic convergence guarantees, while there exist several applications that require finite-time convergence of the estimates.

In the context of finite-time estimation, the sliding-mode methodology is one of the preferred choice for the ease of implementation (see [3], [6] and [7]). However, since the finite-time convergence is attained by high-gain injection, it is very sensitive to measurement noise. Two alternative deadbeat estimation methodologies not requiring high-gain injection, can be found in [8], [9] where integral algebra is used, to reconstruct initial conditions (as presented in [10],

[11]) or eliminate the effect of the initial conditions [9], [12], [13].

Their effectiveness in numerous applications has been shown in [9], [14], [15] and [16]. On the other hand, due to the internal instability issue, periodic resettings are typically needed in these method as in [17] and [16], whose significance has been stated through a simulation comparison in [18]. To overcome this drawback, a kernel-based deadbeat estimation methodology has been proposed recently in [13] and [19] exploiting Volterra operators that allow to avoid the periodic resetting and high-gain injection showing huge potential in many applications (e.g. [20], [21], [22]).

However, most of the finite-time works in the literature address the parameters estimation and the state observation separately, such as [23]. In order to achieve the simultaneous estimation and reduce the complexity, in this work we propose a single estimator for linear MISO systems, jointly providing the estimates of parameters and state variables in arbitrarily short time under persistency of excitation. By suitably shaping the kernel function of the Volterra operator, the terms depending on the initial conditions are suppressed thus guaranteeing the finite-time convergence of the estimates with an internally stable realization and without involving high gain injection.

II. PROBLEM STATEMENT

Consider a MISO system in the input-output form:

$$\mathcal{S}_{u \to y} : \begin{cases} y^{(n)}(t) = \sum_{i=0}^{n-1} a_i y^{(i)}(t) + \sum_{k=0}^{M-1} \sum_{j=0}^{m_k-1} b_{k,j} u_k^{(j)}(t), \\ \qquad\qquad \forall t \in \mathbb{R}_{\geq 0}; \\ y^{(i)}(0) = y_0^{(i)}, \, i \in \{0, \dots, n-1\}; \\ u_k^{(j)}(0) = u_{k,0}^{(j)}, \, j \in \{0, \dots, m_k - 1\} \end{cases} \tag{1}$$

with $n \in \mathbb{Z}_{>0}$, $m_k \in \mathbb{Z}_{>0}$, $m_k \leq n$, $\forall k \in \{0, \dots, M-1\}$. The values of the constant parameters $a_i \in \mathbb{R}$, $b_{k,j} \in \mathbb{R}$ with $i \in \{0, \dots, n-1\}$, $j \in \{0, \dots, m_k-1\}$, $k \in \{0, \dots, M-1\}$ are unknown. The only known elements are the input and output signals, i.e. $u_k(t)$ and $y(t)$, while their derivatives as well as the initial conditions $u_{k,0}^{(j)}(0)$ and $y^{(j)}(0)$ are assumed to be unavailable.

Consider the state-space realization of system (1) in the observer canonical form :

$$\mathcal{S}_{u \to z \to y} : \begin{cases} z^{(1)}(t) = A z(t) + B u(t), \\ y(t) = c^\top z(t), \quad \forall t \in \mathbb{R}_{\geq 0} \end{cases} \tag{2}$$

where $z(t) \triangleq [z_0(t)\, z_1(t)\, \ldots\, z_r(t)\, \ldots z_{n-1}(t)]^\top \in \mathbb{R}^n$ is the state vector and the input vector is defined as $u(t) \triangleq [u_0(t), u_1(t), \ldots, u_{M-1}(t)]$. $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times M}$ and $c \in \mathbb{R}^n$ take on the following forms:

$$A = \begin{bmatrix} a_{n-1} & 1 & 0 & \cdots & 0 \\ a_{n-2} & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ a_1 & 0 & \ldots & 0 & 1 \\ a_0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \; c = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$B = [b_0, b_1, \ldots, b_{M-1}],$$

$b_k = [0, \ldots, 0, b_{k,m_k-1}, \ldots, b_{k,0}]^\top, \forall k \in \{0, \ldots, M-1\}$. The state variables of the realization (2) can be expressed in terms of the input-output derivatives as

$$z_r(t) = y^{(r)}(t) - \sum_{j=0}^{r-1} a_{n-r+j} y^{(j)}(t) - \sum_{k=0}^{M-1} \sum_{j=0}^{r-1+m_k-n} b_{k,n-r+j} u_k^{(j)}(t). \tag{3}$$

where the convention $\sum_{j=0}^k \{\cdot\} = 0$, for $k < 0$ has been used. The goal of this paper is to simultaneously provide deadbeat estimates of the state vector $z(t)$ and of the unknown parameters a_i and $b_{k,j}$ based on the I/O measurements $y(t)$ and $u(t)$. Thanks to the proposed approach, the unknown initial conditions $z_r(0)$, $r \in \{0, \ldots, n-1\}$ do not affect the estimation transient.

III. FINITE-TIME PARAMETER-STATE JOINT ESTIMATION

For readers' convenience and to make the paper sufficiently self-contained, some key facts instrumental to the proposed framework are recalled in Appendix A. Moreover, the interested reader is referred to [13] and [19] for a deeper insight on the algebra of Volterra integral operators.

Consider an n-th order Bivariate Feedthrough Non-asymptotic Kernel (BF-NK proposed in [19]) function $K_h(t, \tau)$ (setting $N \geq n$) in the shape of

$$K_h(t, \tau) = e^{-\omega_h(t-\tau)} \left(1 - e^{-\bar\omega t}\right)^N, \tag{4}$$

tuned by user-defined positive parameters $\omega_h \in \mathbb{R}_{>0}$ and $\bar\omega \in \mathbb{R}_{>0}$. Indeed, this kernel function is characterized by two outstanding features:

- It is non-asymptotic up to the N-th order i.e. $K^{(i)}(t, 0) = 0, \forall i \in \{1 \ldots n-1\}$, which removes the dependence on the initial conditions (see [13]);
- The Volterra operators induced by this kernel can be implemented as an LTV system, processing the available I/O signals and producing the transformed signals as output.

Therefore, the image function of the output's derivatives in the operational domain can be expressed as

$$\left[V_{K_h} y^{(i)}\right](t) \triangleq \int_0^t K_h(t, \tau) y^{(i)}(\tau) d\tau$$
$$= \sum_{j=0}^{i-1} (-1)^{i-j-1} y^{(j)}(t) K_h^{(i-j-1)}(t, t)$$
$$+ (-1)^i \left[V_{K_h^{(i)}} y\right](t), i \in \{1, \ldots, n-1\}. \tag{5}$$

Similarly, the input derivatives transform into:

$$\left[V_{K_h} u_k^{(i)}\right](t) = \sum_{j=0}^{i-1} (-1)^{i-j-1} u_k^{(j)}(t) K_h^{(i-j-1)}(t, t)$$
$$+ (-1)^i \left[V_{K_h^{(i)}} u_k\right](t), i \in \{1, \ldots, m_k - 1\}. \tag{6}$$

Consider the $i = 1$ case. Then (5) becomes

$$[V_{K_h^{(1)}} y](t) = y(t) K_h(t, t) - [V_{K_h} y^{(1)}](t), \tag{7}$$

and replacing the generic y with $y^{(n-1)}$, it holds that

$$[V_{K_h^{(1)}} y^{(n-1)}](t) = y^{(n-1)}(t) K_h(t, t) - [V_{K_h} y^{(n)}], \tag{8}$$

which is equivalent to the following expression according to the I/O model (1) and thanks to the linearity

$$[V_{K_h^{(1)}} y^{(n-1)}](t) = y^{(n-1)}(t) K(t, t) - \sum_{i=0}^{n-1} a_i [V_{K_h} y^{(i)}](t)$$
$$- \sum_{k=0}^{M-1} \sum_{i=0}^{m_k-1} b_{k,i} [V_{K_h} u_k^{(i)}](t). \tag{9}$$

Substituting (5) and (6) into (9), we obtain

$$(-1)^{n-1} [V_{K_h^{(n)}} y](t) + \sum_{j=0}^{n-2} (-1)^{n-2-j} y^{(j)}(t) K_h^{(n-j-1)}(t, t)$$
$$= y^{(n-1)}(t) K_h(t, t) - \sum_{i=0}^{n-1} a_i \Big((-1)^i [V_{K_h^{(i)}} y](t)$$
$$+ \sum_{j=0}^{i-1} (-1)^{i-j-1} y^{(j)}(t) K_h^{(i-j-1)}(t, t) \Big)$$
$$- \sum_{k=0}^{M-1} \sum_{i=0}^{m_k-1} b_{k,i} \Big((-1)^i [V_{K_h^{(i)}} u_k](t)$$
$$+ \sum_{j=0}^{i-1} (-1)^{i-j-1} u_k^{(j)}(t) K_h^{(i-j-1)}(t, t) \Big). \tag{10}$$

After some cumbersome algebra, (10) can be written as

$$(-1)^{n-1} [V_{K_h^{(n)}} y](t) + \sum_{i=0}^{n-1} a_i (-1)^i [V_{K_h^{(i)}} y](t)$$
$$+ \sum_{k=0}^{M-1} \sum_{i=0}^{m_k-1} b_{k,i} (-1)^i [V_{K_h^{(i)}} u_k](t)$$
$$= \sum_{r=0}^{n-1} (-1)^{n-r-1} K_h^{(n-r-1)}(t, t) \Big(y^{(r)}(t)$$
$$- \sum_{j=0}^{r-1} a_{n-r-j} y^{(j)}(t) - \sum_{k=0}^{M-1} \sum_{j=0}^{r-1+m-n} b_{k,n-r-j} u^{(j)}(t) \Big),$$

which, in turn, can be written in the following compact form by recalling (3):

$$(-1)^{n-1} [V_{K_h^{(n)}} y](t) = -\sum_{i=0}^{n-1} a_i (-1)^i [V_{K_h^{(i)}} y](t)$$
$$- \sum_{k=0}^{M-1} \sum_{i=0}^{m_k-1} b_{k,i} (-1)^i [V_{K_h^{(i)}} u_k](t) + \sum_{r=0}^{n-1} \gamma_h(t) z_r(t), \tag{11}$$

where $\gamma_{h,r}(t) = (-1)^{n-r-1} K_h^{(n-r-1)}(t, t)$.

It is worth noting that the right hand side of (11) is linear with respect to the parameters $a_i, b_{k,i}$ and the state variables

$z_r(t)$. Therefore, (11) can be written in the form of a linear constraint

$$(-1)^{n-1}[V_{K_h^{(n)}}y](t) = \nu_h(t)^\top \theta(t), \qquad (12)$$

where

$$\theta(t) \triangleq \Big[a_0, \ldots, a_{n-1}, b_{0,0}, \ldots, b_{0,m_0-1},$$
$$\ldots, b_{M-1,m_{M-1}-1}, z_0(t), \ldots, z_{n-1}(t)\Big].$$

$$\nu_h(t) = \Big[- [V_{K_h}y](t), \ldots, (-1)^n[V_{K_h^{(n-1)}}](t),$$
$$- [V_{K_h}u_0](t), \ldots, (-1)^{m_{M-1}}[V_{K_h^{(m_{M-1}-1)}}u_{M-1}](t),$$
$$\gamma_{h,0}(t), \ldots, \gamma_{h,n-1}(t)\Big]^\top.$$

Note that the vector $\nu_h(t)$ is composed of known functions consisting in the weak derivatives of the kernel $K_h(t, \tau)$ and in the transformed I/O signals. Now we will show that the transformed signals that participate as the elements of $\nu_h(t)$ can be obtained as the output of an LTV dynamic system for any bounded I/O signals $y(t)$ and $u_k(t)$. Let us define the transformed signal vectors

$$\xi_{h,y} \triangleq \Big[[V_{K_h}y](t), \ldots, [V_{K_h^{(n)}}y](t)\Big]$$

$$\xi_{h,u_k} \triangleq \Big[[V_{K_h}u_k](t), \ldots, [V_{K_h^{(m_k-1)}}u_k](t)\Big],$$

with $k \in \{0, \ldots, M-1\}$. Thanks to the fact that the kernel $K_h(t, \tau)$ verifies the identity

$$\frac{\partial}{\partial t}K_h^{(i)}(t, \tau) = -\omega_h K_h^{(i)}(t, \tau), \qquad (13)$$

the auxiliary vectors can be calculated by the following the internally stable LTV system

$$\begin{cases} \xi_{h,\star}^{(1)}(t) = G_{h,\star}\xi_{h,\star}(t) + E_{h,\star}(t) \star (t), \\ \xi_{h,\star}(0) = 0, \end{cases} \qquad (14)$$

where $\star(t)$ represents the variables $y(t)$ and $u_k(t)$, $k \in \{0, \ldots, M-1\}$ and

$$G_{h,y} = diag(-\omega_h) \in \mathbb{R}^{(n+1)\times(n+1)},$$
$$G_{h,u_k} = diag(-\omega_h) \in \mathbb{R}^{m_k \times m_k},$$

$$E_{h,y}(t) = \begin{bmatrix} K_h(t,t) \\ \vdots \\ K_h^{(n)}(t,t) \end{bmatrix}, E_{h,u_k}(t) = \begin{bmatrix} K_h(t,t) \\ \vdots \\ K_h^{(m_k-1)}(t,t) \end{bmatrix}.$$

With the purpose of solving (12) for $n_\delta \triangleq 2n + \sum_{k=0}^{M-1} m_k$ unknown elements (some of them are not constant), the augmentation of the linear constraint (12) is necessary. In this connection, we make use of n_δ kernel functions in the form of (4) with a common $\bar{\omega}$ but different ω_h, $h \in \{0, \ldots, n_\delta-1\}$ to construct n_δ constraints in the form of (12). As a result, the estimation is obtained by solving the following algebraic linear equation for the unknown vector θ:

$$\kappa(t) = \Gamma(t)\theta, \qquad (15)$$

where

$$\kappa(t) = \Big[(-1)^{n-1}[V_{K_0^{(n)}}y](t), \ldots, (-1)^{n-1}[V_{K_{n_\delta-1}^{(n)}}y](t)\Big]^\top,$$
$$\Gamma(t) = [\nu_0(t), \nu_1(t), \ldots, \nu_{n_\delta-1}(t)]^\top.$$

To guarantee the invertibility of the matrix $\Gamma(t)$, the following assumption is needed.

Assumption 1: Given the I/O measurement $y(t)$ and $u_k(t)$ in (1) and the designed kernel (4), there exist some $\epsilon_h \in \mathbb{R}_{>0}$ and $t_{\epsilon,h} \in \mathbb{R}_{>0}$ such that

$$\int_{t-t_{\epsilon,h}}^t \nu_h(\tau)\nu_h(\tau)^\top \geq \epsilon_h I, \forall t > 0, \qquad (16)$$

for all $h \in \{0, \ldots, n_\delta - 1\}$.

Assumption 1 immediately follows from the definition of persistency of excitation [24]:

Definition 3.1: (Persistency of excitation) The function vector $q(t)$ is said to be persistently exciting (PE) if and only if there exist $\epsilon_q \in \mathbb{R}_{>0}$ and $t_q \in \mathbb{R}_{>0}$ such that

$$\int_{t-t_q}^t q(\tau)q(\tau)^\top d\tau > \epsilon_q I, \forall t \geq 0. \qquad (17)$$

We furthermore need the following definition:

Definition 3.2: (Sufficient richness [25]) The signal $u(t)$ is said to be sufficiently rich of order n_p in an arbitrary finite time interval $[t_1, t_2]$ if there does not exist a non-zero vector $p = [p_0, \ldots, p_{n_p-1}]$, verifying

$$\sum_{i=0}^{n_p-1} p_i u^{(i)}(t) = 0, \forall t \in (t_1, t_2). \qquad (18)$$

In the following proposition, we provide some conditions on the I/O signals to satisfy Assumption 1.

Proposition 3.1: (Implication) The PE condition (16) is satisfied if the I/O signals verify the following conditions

- The input signals $u_k(t)$ are sufficiently rich at least of order m_k, $\forall k \in \{0, \ldots, M-1\}$, $\forall t > 0$;
- The output signal $y(t)$ is sufficiently rich at least of order n, $\forall t > 0$;
- The I/O vector $\zeta(t) \triangleq [y(t), u_0(t), \ldots, u_{M-1}(t)]$ is PE.

Therefore, under Assumption 1, stacking n_δ vectors $\nu_h(t)$ induced by different kernel functions $K_h(t, \tau)$, $h \in \{0, \ldots, n_\delta - 1\}$ ensures that the matrix $\Gamma(t)$ is invertible $\forall t > 0$. In turn, the unknown vector θ can be estimated by

$$\hat{\theta}(t) = \Gamma(t)^{-1}\kappa(t), \forall t > 0. \qquad (19)$$

As a result, a *theoretical instantaneous convergence* of the simultaneous parameter-state estimates can be achieved.

IV. NUMERICAL EXAMPLE

In this section, to show the effectiveness of the proposed method, we consider an MISO LTI system with the form:
$y^{(2)}(t) = a_1 y^{(1)}(t) + a_0 y(t) + b_{11}u_0^{(1)}(t) + b_{20}u_1(t)$, (20)
$\forall t \in \mathbb{R}_{\geq 0}$, where the initial conditions and the parameters of the observer canonical form are $z(0) = [1, -3]^\top$ and $a_1 = -1$, $a_0 = -0.3$, $b_{11} = 2$, $b_{20} = 0.5$. The inputs are $u_0(t) = 10\sin(10t)$ and $u_1(t) = \sin(2t) - 0.5$, which are sufficiently rich of order 2 and 3 respectively. As a result, the output is rich of order 9. We note that the I/O vector $[y(t), u_0(t), u_1(t)]$ is PE due to the linear independence among its components.
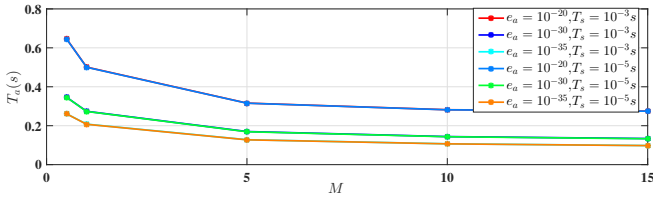
Fig. 1. Effects of activation threshold e_a, kernel parameter M and sampling interval T_s on activation time T_a .
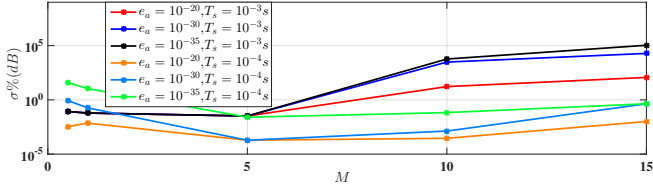

Fig. 2. Effects of activation threshold e_a, kernel parameter M and sampling interval T_s on distortion overshoot $\sigma\%(dB)$.

Fig. 3. Time behavior of the parameter-state joint estimator in noise-free scenario.

A. Noise-free scenario

Firstly, we investigate the performance of the joint estimator in the noise-free scenario. Theoretically, the present estimator is able to achieve instantaneous convergence immediately after $t \neq 0$. However, in the numerical implementation, distortion may occur due to the discretization, significantly depending on the sampling time. Therefore, to avoid the large overshoot caused by inverting a nearly singular matrix in the initial phase, we set an activation threshold e_a to detect the non-singularity of matrix $\Gamma(t)$. The inversion of $\Gamma(t)$ in (19) is only executed when $det(\Gamma(t)) > e_a$. Remarkably, the fast convergence of the estimation is relative to the activation time.

1) Parameters selection: The following parameters need to be set by the designer: the activation threshold e_a, the sampling interval T_s, and the kernel parameters $\omega_h, \bar{\omega}$ in (4). Figures 1 and 2 provide some insight on how the activation time and the overshoot $\sigma\% \triangleq \frac{\text{maximal estimate} - \text{true value}}{\text{true value}} \times 100\%$ are affected by the activation threshold e_a and the sampling interval T_s with different sets of kernels in the form of (4) with $[\omega_0, \ldots, \omega_5] = [1\ 2\ 3\ 4\ 5\ 6] \times M$, $M \in \{0.5, 1, 5, 10, 15\}$ and $\bar{\omega} = 2.5$.

As suggested by intuition, Fig. 1 shows that smaller activation thresholds corresponds to shorter activation times. On the other hand, by increasing the values of the kernel parameters, we obtain shorter activation times, even if the improvement is smaller for larger parameters values. The sampling interval does not have significant effect on the activation time. By analyzing Fig. 2, we see that in general, the overshoot increases with the decrease of the activation threshold. In this regards, designer should consider the trade off between the activation time and the overshoot after activation in order to choose a proper value of e_a. However the overshoot can be remarkably reduced by decreasing the sampling interval, which is reasonable since the distortion is caused by discretization. The overshoot does not have a clear relation with the variation of the kernel parameters, reaching
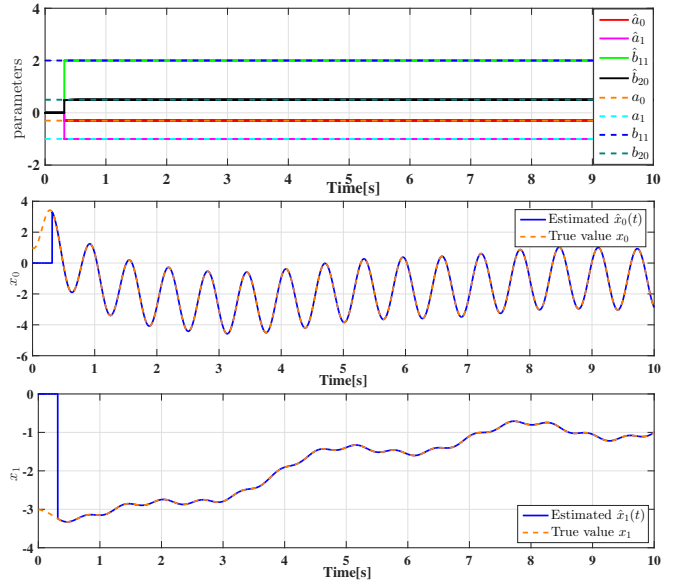
a minimum for M=5.

2) Simulation results: According to this analysis, in this simulation example, to design the proposed parameter-state joint estimator, we use 6 BF-NK functions with $\bar{\omega} = 2.5$, $[\omega_0, \ldots, \omega_5] = [1, 2, 3, 4, 5, 6] \times 5$ and $N = 4$. The activation threshold in chosen as $e_a = 10^{-30}$. The simulations are performed setting the sampling time to $T_s = 10^{-3}s$.

In Fig. 3 the theoretical deadbeat convergence of the proposed method is shown: the estimates of the parameters jump to the true values immediately after activation. At the same time, also the states are accurately reconstructed.

B. Noisy scenario

We examine the estimator performance in a scenario where the measurement $y(t)$ is corrupted by a uniformly distributed random additive noise ranging within $[-0.2, 0.2]$ (see Fig. 4).

1) Parameter selection: We analyze the choice of the parameters in this second case. In order to consider the robustness of the proposed estimator for different parameters configurations, Table. I presents the root mean square error (RMSE) of the parameters and state variables estimates:

$$RMSE(\varsigma) \triangleq \sqrt{\frac{\sum_{n_{rmse}} (e_\varsigma^\top e_\varsigma)}{n_{rmse}}},$$

where ς represents the parameter vector $pars \triangleq [a_0, a_1, b_{11}, b_{20}]^\top$ and the state vector $x(t)$ while e_ς denotes the corresponding estimation error. n_{rmse} represents the number of samples used for RMSE calculation. In order to analyze only the robustness to the noise of the estimators, we do not want to consider the overshoot effect in the calculation of the RMSE, thus in this example we are using the samples from $0.2s$ after activation time until the end of the simulation.

From Table. I, the estimator robustness to the measurement noise decreases as the kernel parameters increase, which

TABLE I
RMSE WITH DIFFERENT PARAMETERS CHOICE

T_s	e_a	M	T_a	RMSE (pars)	RMSE (x)
10^{-3}	10^{-7}	0.5	3.669	0.0876	0.2944
		1	2.027	0.2805	0.3529
		2	1.371	2.0521	0.5572
		3	1.280	15.5913	1.4915
		5	1.381	84.7760	6.0135
	10^{-6}	0.5	4.401	0.0237	1.2348
		1	2.717	0.1066	0.6553
		2	1.825	0.8742	0.2602
		3	1.686	4.9175	0.7505
		5	2.121	25.6606	3.0920
10^{-4}	10^{-7}	0.5	3.6704	0.0215	0.2813
		1	2.0259	0.0860	0.2589
		2	1.3796	0.9318	0.3332
		3	1.2809	5.3712	0.4761
		5	1.3767	44.7607	2.0935
	10^{-6}	0.5	4.4004	0.0089	1.2376
		1	2.7162	0.0405	0.6518
		2	1.8237	0.3483	0.1285
		3	1.6863	1.5522	0.2707
		5	2.1346	9.13733	1.5941

Fig. 4. Measured output $y_d(t)$ in the noise scenario and pure output $y(t)$.

Fig. 5. Time behavior of the parameter-state joint estimator in noisy scenario.

is reasonable because the derivatives of the kernel will be amplified by large parameters, thus increasing the sensitivity of the proposed estimator to the noise. Moreover, larger value of kernel parameters make the modulated noise signal too large to be comparative to the nominal signals, thus giving rise to singularity issues of $\Gamma(t)$. Therefore, the kernel parameters with $M > 10$ will not have enough excitation in presence of noise. Furthermore, we observe in Table. I that the increase of the activation threshold enhances the noise attenuation. However, this is not applicable in the case of state estimation for small kernel parameters ($M \leq 1$), since smaller values of the parameters reduce the determinant of $\Gamma(t)$, and therefore higher activation thresholds make the estimator to be switched off more often. In this regard, designer should consider the tradeoff between activation time, robustness to noise, and persistent excitation, in tuning M and e_a.

2) Simulation results: In this example, we choose $M = 0.5$, $e_a = 10^{-7}$ and $T_s = 10^{-3}s$ obtaining the results depicted in Fig. 5. Please note that the spikes in the noisy estimation are due to numerical reasons: the presence of noise makes the matrix $\Gamma(t)$ close to its singularity. This is likely to happen periodically since the output is sinusoidal signal and the effect of the noise on its derivatives is maximum when the output is at its positive and negative peaks. When $\Gamma(t)$ becomes so that it is not possible to be inverted, the estimator is frozen to a constant value. In digital implementation, we resort to the activation threshold to detect the rank deficiency of $\Gamma(t)$ caused by the presence of noise.
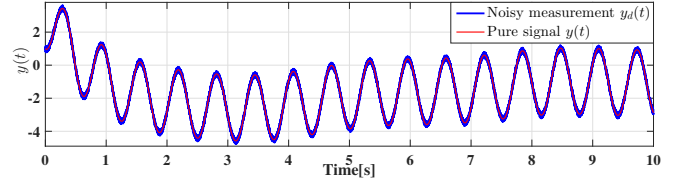
A more comprehensive robustness analysis against measurement disturbance will be presented in future works.

V. CONCLUDING REMARKS

In this paper, a simultaneous parameter-state estimator for MISO linear systems has been developed relying on the kernel-based deadbeat estimation methodology. By making use of suitable Volterra integral operators, the proposed estimator (theoretically) provides exact estimates in the noise-free scenario of both the unknown system parameters and the internal states in finite-time based only on the I/O signals. Remarkably, the transient is instantaneous and independent on the unknown initial conditions. Preliminary simulation experiments are reported showing the effectiveness of the proposed estimator both in the noise-free and noisy scenario.

Future research efforts will be devoted to the case with measurement noise aiming at a comprehensive robustness analysis. Moreover, the extension to some classes of nonlinear systems will be considered as well.

REFERENCES

[1] R. E. Kopp and R. J. Orford, "Linear regression applied to system identification for adaptive control systems," *Aiaa Journal*, vol. 1, no. 10, pp. 2300–2306, 1963.
[2] L. Nelson and E. Stear, "The simultaneous on-line estimation of parameters and states in linear systems," *IEEE Transactions on Automatic Control*, vol. 21, no. 1, pp. 94–98, 1976.

[3] A. B. Proca and A. Keyhani, "Sliding-mode flux observer with online rotor parameter estimation for induction motors," *IEEE Transactions on Industrial Electronics*, vol. 54, no. 2, pp. 716–723, 2007.

[4] A. Concha, L. Alvarez-Icaza, and R. Garrido, "Simultaneous parameter and state estimation of shear buildings," *Mechanical Systems and Signal Processing*, vol. 70, pp. 788–810, 2016.

[5] G. Kreisselmeier, "Adaptive observers with exponential rate of convergence," *IEEE Transactions on Automatic Control*, vol. 22, no. 1, pp. 2–8, 1977.

[6] S. K. Spurgeon, "Sliding mode observers: a survey," *International Journal of Systems Science*, vol. 39, no. 8, pp. 751–764, 2008.

[7] M. Tursini, R. Petrella, and F. Parasiliti, "Adaptive sliding-mode observer for speed-sensorless control of induction motors," *IEEE Transactions on Industry Applications*, vol. 36, no. 5, pp. 1380–1387, 2000.

[8] S. H. Dandach, B. Fidan, S. Dasgupta, and B. D. Anderson, "A continuous time linear adaptive source localization algorithm, robust to persistent drift," *Systems & Control Letters*, vol. 58, no. 1, pp. 7–16, 2009.

[9] M. Mboup, "Parameter estimation for signals described by differential equations," *Applicable Analysis*, vol. 88, no. 1, pp. 29–52, 2009.

[10] G. Rao and H. Unbehauen, "Identification of continuous-time systems," *IEE Proceedings-Control Theory and Applications*, vol. 153, no. 2, pp. 185–220, 2006.

[11] R. E. Kalman, P. L. Falb, and M. A. Arbib, *Topics in mathematical system theory*. McGraw-Hill New York, 1969, vol. 33.

[12] G. Fedele and L. Coluccio, "A recursive scheme for frequency estimation using the modulating functions method," *Applied Mathematics and Computation*, vol. 216, no. 5, pp. 1393–1400, 2010.

[13] G. Pin, A. Assalone, M. Lovera, and T. Parisini, "Non-asymptotic kernel-based parametric estimation of continuous-time linear systems," *IEEE Transaction on Automatic Control*, vol. 61, no. 2, pp. 360–373, 2016.

[14] A. Neves, M. D. Miranda, and M. Mboup, "Algebraic parameter estimation of damped exponentials," in *2007 IEEE European Signal Processing Conference, 2007*. IEEE, 2007, pp. 965–969.

[15] Y. Tian, "Une contribution a l'observation et a l'estimation des systemes linéaires," Ph.D. dissertation, Ecole Centrale de Lille, 2010.

[16] B. Chen, G. Pin, W. Ng, T. Parisini, and S. Hui, "A fast-convergent modulation integral observer for online detection of the fundamental and harmonics in grid-connected power electronics systems," *IEEE Transactions on Power Electronics*, 2016.

[17] H. Sira-Ramirez and M. Fliess, "An algebraic state estimation approach for the recovery of chaotically encrypted messages," *International Journal of Bifurcation and Chaos*, vol. 16, no. 02, pp. 295–309, 2006.

[18] J. Reger, P. Mai, and H. Sira-Ramirez, "Robust algebraic state estimation of chaotic systems," in *2006 IEEE International Conference on Control Applications*, Oct 2006, pp. 326–331.

[19] G. Pin, M. Lovera, A. Assalone, and T. Parisini, "Kernel-based non-asymptotic state estimation for linear continuous-time system," in *2013 IEEE American Control Conference*, Washington, DC, 2013, pp. 3123–3128.

[20] P. Li, G. Pin, G. Fedele, and T. Parisini, "Deadbeat source localization from range-only measurements: A robust kernel-based approach," *IEEE Transactions on Control Systems Technology*, vol. PP, no. 99, pp. 1–11, 2018.

[21] G. Pin, B. Chen, and T. Parisini, "Robust finite-time estimation of biased sinusoidal signals: A volterra operators approach," *Automatica*, vol. 77, pp. 120–132, 2017.

[22] P. Li, F. Boem, G. Pin, and T. Parisini, "Distributed fault detection and isolation for interconnected systems: a non-asymptotic kernel-based approach," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 1013–1018, 2017, 20th IFAC World Congress. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2405896317305219

[23] J. Jouffroy and J. Reger, "Finite-time simultaneous parameter and state estimation using modulating functions," in *2015 IEEE Conference on Control Applications*. IEEE, 2015, pp. 394–399.

[24] N. Shimkin and A. Feuer, "Persistency of excitation in continuous-time systems," *Systems & Control Letters*, vol. 9, no. 3, pp. 225–233, 1987.

[25] G. Kreisselmeier and G. Rietze-Augst, "Richness and excitation on an interval-with application to continuous-time adaptive control," *IEEE Transactions on Automatic Control*, vol. 35, no. 2, pp. 165–171, 1990.

APPENDIX

A. Non-asymptotic Volterra integral operator algebra

Consider a signal $x(t) \in \mathbb{R}$, $\forall t \geq 0$ having its i-th weak derivative denoted by $x^{(i)}(t)$. Moreover, given a bivariate kernel function $K(\cdot, \cdot)$, its i-th order weak derivative with respect to the second argument will be denoted as $K^{(i)}(t, \tau)$, $i \in \mathbb{Z}_{\geq 0}$.

The Volterra transformations act on the Hilbert space $\mathcal{L}^2_{loc}(\mathbb{R}_{\geq 0})$ of locally square-integrable functions with domain $\mathbb{R}_{\geq 0}$ and range $\mathbb{R}$ (*i.e.*, $x(\cdot) \in \mathcal{L}^2_{loc}(\mathbb{R}_{\geq 0}) \Leftrightarrow (x(\cdot) : \mathbb{R}_{\geq 0} \to \mathbb{R}) \wedge (\int_B |x(t)|^2 dt < \infty$, $\forall$ compact $B \subset \mathbb{R}_{\geq 0})$. Given a function $x \in \mathcal{L}^2_{loc}(\mathbb{R}_{\geq 0})$, its image through the Volterra (linear, integral) operator V_K induced by a Hilbert-Schmidt $\mathcal{HS}$ *Kernel Function* $K(\cdot, \cdot) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is usually denoted by $[V_K x](\cdot)$, and is defined by the inner product:

$$[V_K x](t) \triangleq \int_0^t K(t, \tau) x(\tau) d\tau, \ t \in \mathbb{R}_{\geq 0}.$$

Moreover, the transformed signal $[V_K x](t)$, for $t \geq 0$, can be obtained as the output of a dynamic system described by the following scalar integro-differential equation:

$$\begin{cases} \xi^{(1)}(t) &= K(t,t)x(t) + \int_0^t \left(\frac{\partial}{\partial t}K(t,\tau)\right) x(\tau) d\tau \\ [V_K x](t) &= \xi(t) \end{cases} \tag{21}$$

where $\xi(0) = \int_0^0 K(0,\tau)x(\tau) = 0 \ d\tau$ and $\xi^{(1)}(0) = 0$.

It is worth noting that the system (21) is internally stable as long as the kernel function and its weak derivatives w.r.t. the second argument are bounded.

Lemma 1.1: [13] For a given $i \geq 0$, consider a signal $x(\cdot) \in \mathcal{L}^2(\mathbb{R}_{\geq 0})$ that admits a i-th weak derivative in $\mathbb{R}_{\geq 0}$ and a kernel function $K(\cdot, \cdot) \in \mathcal{HS}$, admitting the i-th derivative (in the conventional sense) with respect to the second argument. Then, it holds that:

$$\left[V_K x^{(i)}\right](t) = \sum_{j=0}^{i-1} (-1)^{i-j-1} x^{(j)}(t) K^{(i-j-1)}(t,t)$$
$$+ \sum_{j=0}^{i-1} (-1)^{i-j} x^{(j)}(0) K^{(i-j-1)}(t,0) + (-1)^i \left[V_{K^{(i)}} x\right](t) \tag{22}$$

that is, the function $\left[V_K x^{(i)}\right](\cdot)$ is non-anticipative with respect to the lower-order derivatives $x(\cdot)$, $x^{(1)}(\cdot), \ldots, x^{(i-1)}(\cdot)$. $\square$

The properties of the Volterra operator depend significantly on the shape of the kernel function. The non-asymptotic kernel functions defined in the following lines play a significant role in deadbeat estimation[19], [13].

Definition 1.1: [13] If a kernel $K(\cdot, \cdot) \in \mathcal{HS}$ which is at least $(i-1)$-th order differentiable with respect to the second argument, verifies the condition

$$K^{(j)}(t,0) = 0, \ \forall j \in \{0, 1, \ldots, i-1\} \tag{23}$$

then, it is called an *i-th order non-asymptotic kernel*.