2017 IEEE Conference on Control Technology and Applications (CCTA)
August 27-30, 2017. Kohala Coast, Hawai'i, USA

Double-Sided Kernel Observer for Linear Time-Varying Systems*

Debarshi Patanjali Ghoshal[1], Hannah Michalska[2]

***Abstract*— A double-sided input-output kernel functional representation is developed for the class of totally observable linear time-varying systems with inputs. The double-sided kernel representation is immediately applicable as part of a non-asymptotic state observer for observable LTV systems. In the absence of output measurement noise the observer provides exact state values of the system state in arbitrarily short time. It also shares the usual superior features of algebraic observers such as independence of the initial conditions of the system and good noise attenuation properties. Other advantages of the double-sided input-output kernel functional representation of linear systems are elucidated as the concept can be employed to construct state and parameter estimators for flat nonlinear systems.**

I. INTRODUCTION

The concept of non-asymptotic state and parameter estimation is now gaining interest as the development of more efficient control technologies poses new challenges. Hybrid or switched controllers are best matched with dead-beat observers to deliver optimal closed loop performance. The recursive estimation algorithms, with the family of Kalman filters as their prominent representative, admittedly feature powerful noise rejection properties. However, there are also well known problems with their implementation and tuning, as any errors in the assumptions about the system model, noise characteristic, or *initial conditions of the system* can have destabilizing effect on the filter resulting in its slow convergence.

On the other hand, the concept of flatness introduced by Fliess [1], [2], clearly calls for repeated system output differentiation as the principal tool for construction of dead-beat, algebraic observers. The algebraic observers would be an obvious choice if not that accurate higher order differentiation of noisy signals is considered impractical. Generally, this would seem true if no deterministic behavioral property of the system can be assumed. If, however, one can infer the existence of a differential invariant of the system that holds regardless of noise, the situation changes dramatically. A differential invariant is a function $\mathcal{J}(t, y(t), y^{(1)}(t), ..., y^{n}(t))$ of time, output, and a number of its derivatives which is constant under the action of the flow of the system. Such a deterministic quantity is inherent in all linear systems as their system matrix satisfies the Cayley-Hamilton equation that results in an equation that binds the output and its derivatives. The validity of the invariant can then be enforced during signal differentiation resulting in considerable reduction of uncertainty such as measurement noise; see [3].

In the worst case, when no information about the dynamics of the signal-generating system can be made, one can assume an "approximate, plausible model" e.g. that a sufficiently high derivative of the signal vanishes identically; i.e. that $y^{(n)} \equiv 0$. Such an approach has been adopted for differentiation of arbitrary smooth signals in many works of Fliess [1], [2], [4], [5], Sira-Ramirez [6], Mboup [7], [8], et al., and is equivalent to the assumption that the signal to be differentiated has a finite Taylor series expansion. These and similar approaches have been explored for the purpose of state and parameter estimation in both linear and nonlinear flat systems in several works, albeit not referring explicitly to differential invariance; see Fliess [1], [2], [4], [5], Sira-Ramirez [6]. The algebraic estimation approaches developed there have one problem in common which is that of a singularity occuring at the beginning of the estimation horizon. The singularity is introduced by the way in which the operation of differentiation is substituted by integration (necessitating the presence of the term $1/t$ at $t = 0$), as well as the adopted domain of the conversion that is assumed to be the half-line $[0, \infty)$; see [6]. It is argued that such singularity is not critical as the correct estimates are achieved away from zero, i.e. on some subinterval (ϵ, ∞). However, the value of $\epsilon > 0$ is case-dependent and difficult to estimate beforehand which goes against the claim that estimation can be carried out in arbitrary short time. More importantly, the singulaties re-occur each time the estimation process needs to be re-started due to accumulation of numerical errors of integration.

The work presented here is an extension of the results in [3] that pertained only to linear time invariant systems without exogenous input, but in which the singularity problem was successfully removed thus truly permitting uniform estimation precision on arbitrary horizons. The latter was achieved by seeking a double-sided kernel functional representation of the differential invariant involving the system output on a finite rather than infinite horizon $[t_a, t_b] \subset \mathbb{R}^+$. A similar approach is adopted here to derive a double-sided kernel representation of the input-output behavior of totally observable linear time-varying systems with exogenuous inputs. The novel non-singular system representation, together with a recursive non-singular kernel-based formula for computation of output derivatives immediately yields a non-asymptotic state observer, much like the one presented in [9], but of superior performance. The estimation window $[t_a, t_b]$ is dragged forward as new output measurements arrive, which can be carried out safely in the absence of

*This work was supported by The National Science & Engineering Research Council of Canada

[1] Authors are with the Department of Electrical, Computer & Software Engineering, McGill University, 3480 University Street, Montreal H3A 2A7, Quebec, Canada.

[2] Corresponding author hannah.michalska@mcgill.ca

singularities. The observer presented here does not need the notion of "initial conditions" as the latter have been removed from the "behavioral" kernel system representation. The measurement noise is "filtered naturally" as all estimates are obtained as output functions of integral operators.

Unlike in [9], the derivation of the estimator equations is carried out directly in the time domain relying on the *controlled system invariance* constituted by the original differential input-output relation in the system. The constructed dead-beat observer is remarkably robust with respect to Gaussian measurement noise. This is due to the averaging effect of iterated integration as explained by Fliess in [10].

The properties and advantages of the introduced kernel representation are discussed and its relevance to further applications is explained.

II. A KERNEL REPRESENTATION OF AN LTV SYSTEM WITH EXOGENOUS INPUT

A. Model Assumptions

The LTV system considered is assumed to be stated in the state space form:

$$\dot{x} = A(t)x + B(t)u \ ; \quad y = C(t)x \tag{1}$$

where the state, input and output are: $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^d$; with $A(t) \in \mathbb{R}^{n\times n}$, $B(t) \in \mathbb{R}^{n\times m}$, and $C(t) \in \mathbb{R}^{d\times n}$ are known, $n-1$ times differentiable matrix functions of time.

The following definitions and criteria for complete and total observability of a general system (1) can be found in [11] and [12]:

Definition 1

- System (1) is completely observable on a time interval $[t_0, t_f]$ if any initial state can be determined from the knowledge of the system output and input, $y(t)$ and $u(t)$, on the interval $t \in [t_0, t_f]$.
- System (1) is totally observable on a time interval $[t_0, t_f]$ if it is completely observable on any subinterval of $[t_0, t_f]$.

Theorem 1: [12] System (1) restricted to the time interval $[t_0, t_f]$ is:

- completely observable if $rank\ \mathcal{O}(t) = n$ for $t \in [t_0, t_f]$;
- totally observable if and only if $rank\ \mathcal{O}(t) = n$ on any subinterval of $[t_0, t_f]$;

where the observability matrix is defined by

$$\begin{aligned} &\mathcal{O} = \{S_0(t), S_1(t), \ldots, S_{n-1}(t)\} \qquad (2)\\ &S_0(t) = C(t)^T;\\ &S_{k+1}(t) = A(t)^T S_k(t) + \dot{S}_k(t); k = 0, \ldots, n-2 \end{aligned}$$

B. A Controlled Differential Invariant - the Input-Output Equation

Henceforth, only single-input, single-output and strictly proper systems will be considered; i.e. $m = 1$, $m < n$, with $d = 1$ so $u, y \in \mathbb{R}$.

With assumption of total observability, the state variable of system (1) can be expressed in terms of the system input, output, and their derivatives; see e.g. [9] for a simple proof of this fact:

Theorem 2: [9] Under the total observability assumption the input-output relation in system (1) is of the form:

$$\sum_{i=0}^{n} a_i(t)y^{(i)} + \sum_{i=0}^{m} b_i(t)u^{(i)} = 0 \quad \text{ for } t \in [t_0, t_f] \tag{3}$$

with $a_n = 1$, $m < n$. The state x can be expressed as a linear function of the input and output and their derivatives:

$$x(t) = \begin{pmatrix} \Gamma_0 \\ \Gamma_1 \\ \Gamma_2 \\ \vdots \\ \Gamma_{n-1} \end{pmatrix}^{-1} \left[\begin{pmatrix} y \\ y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(n-1)} \end{pmatrix} - M \begin{pmatrix} u \\ u^{(1)} \\ u^{(2)} \\ \vdots \\ u^{(n-2)} \end{pmatrix} \right] \tag{4}$$

where

$$\Gamma_0(t) = C(t)$$

$$\Gamma_k(t) = \left(\left(A(t)^T + \frac{d}{dt} \right)^k C(t)^T \right)^T \ ; \quad 0 < k < n$$

$$M = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 \\ \Delta_{11} & 0 & 0 & \ldots & 0 \\ \Delta_{21} & \Delta_{22} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Delta_{(n-1)1} & \Delta_{(n-1)2} & \Delta_{(n-1)3} & \ldots & \Delta_{(n-1)(n-1)} \end{pmatrix}$$

$$\Delta_{k0}(t) = \Gamma_k(t)B(t)$$

$$\Delta_{kj}(t) = \begin{cases} C(t)B(t); & \text{if } j = k \\ \Delta_{(k-1)(j-1)}(t) + \frac{d}{dt}\Delta_{(k-1)j}(t); & \text{if } 1 \le j < k \end{cases} \tag{5}$$

The invertibility of the matrix in (4) is guaranteed by total observability assumption on the system; see the definition of the observability matrix in Theorem 1.

It should be noted that equation (3) in fact constitutes a "controlled differential invariant" of system (1) which will next be used to derive an alternative non-singular integral kernel representation of the system.

C. Derivation of the Double-Sided Kernel Representation of the System Input-Output Relation

Given an arbitrary time interval $[a, b]$, referred to as the estimation window, it is possible to develop a surrogate integral-equation model for (1). Such an embedding also delivers kernel integral formulae for computation of all the derivatives of the system output over $[a, b]$. The following result is valid:

Theorem 3: There exist Hilbert-Schmidt kernels K_y, K_u, K_y^i, K_u^i, $i = 1, \cdots n-1$, and functions f_y^i, f_u^i, $i = 0, \cdots n-2$, defined respectively on $[a, b] \times [a, b]$ and $[a, b]$ such that the output y of system (1) satisfies the following integral equation on $[a, b]$

$$y(t) = \int_a^b K_y(t,\tau)y(\tau)\ d\tau + \int_a^b K_u(t,\tau)u(\tau)\ d\tau \tag{6}$$

while the derivatives of the output $y^{(1)}, \cdots y^{(n-1)}$ for $i = 1, \cdots n-1$ satisfy the recursive relationships :

$$y^{(i)}(t) = \sum_{k=0}^{i-1} f_y^k(t) y^{(k)}(t) + \sum_{k=0}^{i-1} f_u^k(t) u^{(k)}(t) + \int_a^b K_y^i(t,\tau) y(\tau)\, d\tau + \int_a^b K_u^i(t,\tau) u(\tau)\, d\tau \quad (7)$$

□

Hilbert-Schmidt kernels are square integrable functions on $L^2[a,b] \times L^2[a,b]$.

While the analytical formulae for the kernels for a general n-dimensional system are notationally complex, their derivation is surprisingly straightforward. For reasons of brevity and complete lucidity of exposition, the proof of the theorem is conducted here for a system of order two. The procedure is carried out entirely in the time domain and extends immediately to an arbitrary dimension n. The general formulae can then be obtained by induction.

Proof for $n = 2$:

To this end, let $n = 2$, and consider the input output equation on $[a,b]$.

$$y^{(2)}(t) + a_1(t) y^{(1)}(t) + a_0(t) y(t) + b_1(t) u^{(1)}(t) + b_0(t) u(t) = 0 \quad (8)$$

The general idea leading to the desired result of Theorem 3 is to lower the order of the output derivatives in (8) (until no derivatives appear) while simultaneously removing the influence of any pre-existing initial conditions for (8). The initial conditions are removed by pre-multiplication by suitable "annihilator functions" that vanish, together with their derivatives of order up to $n-1$, at the endpoints of $[a,b]$. The process is carried out as follows.

Pre-multiplying (8) by the terms $(\xi - a)^2$ and $(b - \zeta)^2$ (annihilators at $\xi = a$ and $\zeta = b$), respectively, gives

$$(\xi-a)^2 y^{(2)}(\xi) + (\xi-a)^2 a_1(\xi) y^{(1)}(\xi) + (\xi-a)^2 a_0(\xi) y(\xi) + (\xi-a)^2 b_1(\xi) u^{(1)}(\xi) + (\xi-a)^2 b_0(\xi) u(\xi) = 0 \quad (9)$$

$$(b-\zeta)^2 y^{(2)}(\zeta) + (b-\zeta)^2 a_1(\zeta) y^{(1)}(\zeta) + (b-\zeta)^2 a_0(\zeta) y(\zeta) + (b-\zeta)^2 b_1(\zeta) u^{(1)}(\zeta) + (b-\zeta)^2 b_0(\zeta) u(\zeta) = 0 \quad (10)$$

Double integration of the terms in (9) over an interval $[a, a+\tau]$ while repeatedly using integration by parts to lower the order of output derivatives appearing in the integrands, yields the following results when proceeding term by term, (e.g. 1i2 stands for "first term, second iterated integral"):

$$\text{1i1} \int_a^{a+\tau} (\xi-a)^2 y^{(2)}(\xi) d\xi = \tau^2 y^{(1)}(a+\tau) - 2\tau y(a+\tau) + \int_a^{a+\tau} 2y(\xi) d\xi \quad (11)$$

$$\text{1i2} \int_a^{a+\tau} \int_a^{\psi} (\xi-a)^2 y^{(2)}(\xi) d\xi d\psi = \tau^2 y(a+\tau) - 4\int_a^{a+\tau} (\psi-a) y(\psi) d\psi + 2\int_a^{a+\tau} \int_a^{\psi} y(\xi) d\xi d\psi \quad (12)$$

$$\text{2i1} \int_a^{a+\tau} (\xi-a)^2 a_1(\xi) y^{(1)}(\xi) d\xi = \tau^2 a_1(a+\tau) y(a+\tau) - \int_a^{a+\tau} \{2(\xi-a)a_1(\xi) + (\xi-a)^2 a_1^{(1)}(\xi)\} y(\xi) d\xi \quad (13)$$

$$\text{2i2} \int_a^{a+\tau} \int_a^{\psi} (\xi-a)^2 a_1(\xi) y^{(1)}(\xi) d\xi d\psi = \int_a^{a+\tau} (\psi-a)^2 a_1(\psi) y(\psi) d\psi - \int_a^{a+\tau} \int_a^{\psi} \{2(\xi-a)a_1(\xi) + (\xi-a)^2 a_1^{(1)}(\xi)\} y(\xi) d\xi d\psi \quad (14)$$

$$\text{3i1} \int_a^{a+\tau} (\xi-a)^2 a_0(\xi) y(\xi) d\xi \quad (15)$$

$$\text{3i2} \int_a^{a+\tau} \int_a^{\psi} (\xi-a)^2 a_0(\xi) y(\xi) d\xi d\psi \quad (16)$$

$$\text{4i1} \int_a^{a+\tau} (\xi-a)^2 b_1(\xi) u^{(1)}(\xi) d\xi = \tau^2 b_1(a+\tau) u(a+\tau) - \int_a^{a+\tau} \{2(\xi-a)b_1(\xi) + (\xi-a)^2 b_1^{(1)}(\xi)\} u(\xi) d\xi \quad (17)$$

$$\text{4i2} \int_a^{a+\tau} \int_a^{\psi} (\xi-a)^2 b_1(\xi) u^{(1)}(\xi) d\xi d\psi = \int_a^{a+\tau} (\psi-a)^2 b_1(\psi) u(\psi) d\psi - \int_a^{a+\tau} \int_a^{\psi} \{2(\xi-a)b_1(\xi) + (\xi-a)^2 b_1^{(1)}(\xi)\} u(\xi) d\xi d\psi \quad (18)$$

$$\text{5i1} \int_a^{a+\tau} (\xi-a)^2 b_0(\xi) u(\xi) d\xi \quad (19)$$

$$\text{5i2} \int_a^{a+\tau} \int_a^{\psi} (\xi-a)^2 b_0(\xi) u(\xi) d\xi d\psi \quad (20)$$

Collecting the second integrals in (12), (14), (16), (18), (20) while applying the Cauchy formula for multiple iterated integrals and letting $\tau + a \to t$, produces the following equation

$$(t-a)^2 y(t) = \int_a^t K_{Fy}(t,\xi) y(\xi) d\xi + \int_a^t K_{Fu}(t,\xi) u(\xi) d\xi \quad (21)$$

where the kernel functions K_{Fy} and K_{Fu} have the expressions

$$K_{Fy}(t,\xi) = 4(\xi-a) - (\xi-a)^2 a_1(\xi) + (t-\xi)\{-2 + 2(\xi-a)a_1(\xi) + (\xi-a)^2 a_1^{(1)}(\xi) - (\xi-a)^2 a_0(\xi)\} \quad (22)$$

$$K_{Fu}(t,\xi) = -(\xi-a)^2 b_1(\xi) + (t-\xi)\{2(\xi-a)b_1(\xi) + (\xi-a)^2 b_1^{(1)}(\xi) - (\xi-a)^2 b_0(\xi)\} \quad (23)$$

Proceeding in an identical manner with the second "backward" equation (10) while integrating twice over the interval $[b-\sigma, b]$ yields the following expressions for the first and second integrals:

$$1i1 \int_{b-\sigma}^{b}(b-\zeta)^2 y^{(2)}(\zeta)d\zeta = -\sigma^2 y^{(1)}(b-\sigma) - 2\sigma y(b-\sigma) + \int_{b-\sigma}^{b} 2y(\zeta)d\zeta \tag{24}$$

$$1i2 \int_{b-\sigma}^{b}\int_{\phi}^{b}(b-\zeta)^2 y^{(2)}(\zeta)d\zeta d\phi = \sigma^2 y(b-\sigma) - \int_{b-\sigma}^{b} 4(b-\phi)y(\phi)d\phi + \int_{b-\sigma}^{b}\int_{\phi}^{b} 2y(\zeta)d\zeta d\phi \tag{25}$$

$$2i1 \int_{b-\sigma}^{b}(b-\zeta)^2 a_1(\zeta)y^{(1)}(\zeta)d\xi = -\sigma^2 a_1(b-\sigma)y(b-\sigma) + \int_{b-\sigma}^{b}\{2(b-\zeta)a_1(\zeta) - (b-\zeta)^2 a_1^{(1)}(\zeta)\}y(\zeta)d\zeta \tag{26}$$

$$2i2 \int_{b-\sigma}^{b}\int_{\phi}^{b}(b-\zeta)^2 a_1(\zeta)y^{(1)}(\zeta)d\zeta d\phi = -\int_{b-\sigma}^{b}(b-\phi)^2 a_1(\phi)y(\phi)d\phi + \int_{b-\sigma}^{b}\int_{\phi}^{b}\{2(b-\zeta)a_1(\zeta) - (b-\zeta)^2 a_1^{(1)}(\zeta)\}y(\zeta)d\zeta d\phi \tag{27}$$

$$3i1 \int_{b-\sigma}^{b}(b-\zeta)^2 a_0(\zeta)y(\zeta)d\zeta \tag{28}$$

$$3i2 \int_{b-\sigma}^{b}\int_{\phi}^{b}(b-\zeta)^2 a_0(\zeta)y(\zeta)d\zeta d\phi \tag{29}$$

$$4i1 \int_{b-\sigma}^{b}(b-\zeta)^2 b_1(\zeta)u^{(1)}(\zeta)d\zeta = -\sigma^2 b_1(b-\sigma)u(b-\sigma) + \int_{b-\sigma}^{b}\{2(b-\zeta)b_1(\zeta) - (b-\zeta)^2 b_1^{(1)}(\zeta)\}u(\zeta)d\zeta \tag{30}$$

$$4i2 \int_{b-\sigma}^{b}\int_{\phi}^{b}(b-\zeta)^2 b_1(\zeta)u^{(1)}(\zeta)d\zeta d\phi = -\int_{b-\sigma}^{b}(b-\phi)^2 b_1(\phi)u(\phi)d\phi + \int_{b-\sigma}^{b}\int_{\phi}^{b}\{2(b-\zeta)b_1(\zeta) - (b-\zeta)^2 b_1^{(1)}(\zeta)\}u(\zeta)d\zeta d\phi \tag{31}$$

$$5i1 \int_{b-\sigma}^{b}(b-\zeta)^2 b_0(\zeta)u(\zeta)d\zeta \tag{32}$$

$$5i2 \int_{b-\sigma}^{b}\int_{\phi}^{b}(b-\zeta)^2 b_0(\zeta)u(\zeta)d\zeta d\phi \tag{33}$$

As before, collecting the terms in (25), (27), (29), (31), (33), while applying the Cauchy formula for multiple iterated integrals and letting $b-\sigma \to t$, produces the following equation

$$(b-t)^2 y(t) = \int_{t}^{b} K_{By}(t,\zeta)y(\zeta)d\zeta + \int_{t}^{b} K_{Bu}(t,\zeta)u(\zeta)d\zeta \tag{34}$$

where the kernel functions K_{By} and K_{Bu} are

$$K_{By}(t,\xi) = 4(b-\zeta) + (b-\zeta)^2 a_1(\zeta) + (t-\zeta)\{2 + 2(b-\zeta)a_1(\zeta) - (b-\zeta)^2 a_1^{(1)}(\zeta) + (b-\zeta)^2 a_0(\zeta)\} \tag{35}$$

$$K_{Bu}(t,\xi) = (b-\zeta)^2 b_1(\zeta) + (t-\zeta)\{2(b-\zeta)b_1(\zeta) - (b-\zeta)^2 b_1^{(1)}(\zeta) + (b-\zeta)^2 b_0(\zeta)\} \tag{36}$$

Adding equations (21) and (34) side by side while dividing both sides by $C_{[a,b]}(t) = [(t-a)^2 + (b-t)^2]$ yields the desired result (6) with the kernels given by

$$K_y(t,\tau) = \frac{1}{C_{[a,b]}(t)}\begin{cases} K_{Fy}(t,\tau) & \text{for } \tau \le t \\ K_{By}(t,\tau) & \text{for } \tau > t \end{cases} \tag{37}$$

$$K_u(t,\tau) = \frac{1}{C_{[a,b]}(t)}\begin{cases} K_{Fu}(t,\tau) & \text{for } \tau \le t \\ K_{Bu}(t,\tau) & \text{for } \tau > t \end{cases} \tag{38}$$

The kernels for the derivative expression is obtained similarly. Collecting the first integrals in (11), (13), (15), (17) , (19) and letting $\tau + a \to t$, produces the following equation

$$(t-a)^2 y^{(1)}(t) = [2(t-a) - (t-a)^2 a_1(t)]y(t) - (t-a)^2 b_1(t)u(t) + \int_{a}^{t} K_{Fy}^1(t,\xi)y(\xi)d\xi + \int_{a}^{t} K_{Fu}^1(t,\xi)u(\xi)d\xi \tag{39}$$

where the kernel functions K_{Fy}^1 and K_{Fu}^1 are:

$$K_{Fy}^1(\xi) = -2 + 2(\xi-a)a_1(\xi) + (\xi-a)^2 a_1^{(1)}(\xi) - (\xi-a)^2 a_0(\xi) \tag{40}$$

$$K_{Fu}^1(\xi) = 2(\xi-a)b_1(\xi) + (\xi-a)b_1^{(1)}(\xi) - (\xi-a)^2 b_0(\xi) \tag{41}$$

Note that, in general, the kernel functions K_{Fy}^i, K_{Fu}^i, K_{By}^i and K_{Bu}^i would have terms involving t, except when $i = n-1$. Collecting the terms in (24), (26), (28), (30), (32), and letting $b-\sigma \to t$, produces the following equation

$$(b-t)^2 y^{(1)}(t) = -[2(b-t) + (b-t)^2 a_1(t)]y(t) - (b-t)^2 b_1(t)u(t) + \int_{t}^{b} K_{By}^1(t,\zeta)y(\zeta)d\zeta + \int_{t}^{b} K_{Bu}^1(t,\zeta)u(\zeta)d\zeta \tag{42}$$

where the kernel functions K_{By}^1 and K_{Bu}^1 are

$$K_{By}^1(\zeta) = 2 + 2(b-\zeta)a_1(\zeta) - (b-\zeta)^2 a_1^{(1)}(\zeta) + (b-\zeta)^2 a_0(\zeta) \tag{43}$$

$$K_{Bu}^1(\zeta) = 2(b-\zeta)b_1(\zeta) - (b-\zeta)^2 b_1^{(1)}(\zeta) + (b-\zeta)^2 b_0(\zeta) \tag{44}$$

Adding equations (39) and (42) side by side while dividing both sides by $C_{[a,b]}(t)$ yields the desired result (7) with :

$$f_y^0(t) = 2[2t-a-b] + [(b-t)^2 - (t-a)^2]a_1(t) \tag{45}$$

$$f_u^0(t) = -[(b-t)^2 + (t-a)^2]b_1(t) \tag{46}$$

$$K_y^1(t,\tau) = \frac{1}{C_{[a,b]}(t)}\begin{cases} K_{Fy}^1(t,\tau) & \text{for } \tau \le t \\ K_{By}^1(t,\tau) & \text{for } \tau > t \end{cases} \tag{47}$$

$$K_u^1(t,\tau) = \frac{1}{C_{[a,b]}(t)} \begin{cases} K_{Fu}^1(t,\tau) & \text{for } \tau \leq t \\ K_{Bu}^1(t,\tau) & \text{for } \tau > t \end{cases} \tag{48}$$

It should now be clear that the formulae (6) and (7) for a general $n > 2$ can be obtained by pre-multiplication of (3) by the annihilators $h_a(t) := (t-a)^n$ and $h_b(t) := (b-t)^n$, and subsequent n-fold integration over the respective subintervals $[a,t]$ and $[t,b]$ to remove the derivatives of the output to a desired order.

Remarks

• The kernels in Theorem 3 do not have any singularities as the annihilator functions are both positive, strictly monotonic, with their respective ranges: $\mathcal{R}(h_a) = [0, b-a]$ and $\mathcal{R}(h_b) = [b-a, 0]$ for $t \in [a,b]$ implying $C_{[a,b]}(t) > 0$ for $t \in [a,b]$. Implied is the fact that the state estimates will be uniformly accurate over the entire estimation window $[a,b]$. If the input u is piecewise continuous then (7) is still valid almost everywhere on $[a,b]$.

• The equivalent integral system representation (6) acts as a smoother of the measured output signal as the right hand side effectively acts like an integral filter on y. The representation is not unique as different annihilators h_a and h_b can be used as long as they have the obvious properties: $h_a(a) = h_a^{(1)}(a) = \ldots = h_a^{(n-1)}(a) = 0$, and $h_b(b) = h_b^{(1)}(b) = \ldots = h_b^{(n-1)}(b) = 0$. The choice of annihilators should secure improved noise attenuation in cases when the noise characteristics is known.

• The system representation (6) is valid on an arbitrary time window $[a,b]$ and is totally independent of the initial conditions of system (1). Equation (6) is an equivalent system representation in the sense that: if for any given function $u : [a,b] \to \mathbb{R}$ there exists a function $y : [a,b] \to \mathbb{R}$ that satisfies the integral system representation (6), then there exist initial conditions for system (1) such that y is the output in response to input u on $[a,b]$.

III. A NON-ASYMPTOTIC OBSERVER FOR LTV SYSTEMS

The results in Theorem 3 readily deliver a state estimator for system (1) on an arbitrary window $[a,b]$. Given a measured output $y_M : [a,b] \to \mathbb{R}$ in response to a known system input $u_K : [a,b] \to \mathbb{R}$ the estimator equations are:

$$\begin{aligned} y_E(t) &= \int_a^b K_y(t,\tau) y_M(\tau)\, d\tau + \int_a^b K_u(t,\tau) u_K(\tau)\, d\tau \\ y_E^{(i)}(t) &= \sum_{k=0}^{i-1} f_y^k(t) y_E^{(k)}(t) + \sum_{k=0}^{i-1} f_u^k(t) u_K^{(k)}(t) \\ &+ \int_a^b K_y^i(t,\tau) y_M(\tau)\, d\tau + \int_a^b K_u^i(t,\tau) u_K(\tau)\, d\tau \\ x_E(t) &= \begin{pmatrix} \Gamma_0 \\ \vdots \\ \Gamma_{n-1} \end{pmatrix}^{-1} \left[\begin{pmatrix} y_E(t) \\ \vdots \\ y_E^{(n-1)}(t) \end{pmatrix} - M \begin{pmatrix} u_K(t) \\ \vdots \\ u_K^{(n-2)}(t) \end{pmatrix} \right] \end{aligned} \tag{49}$$

where the input derivatives : $u_K^{(k)} : [a,b] \to \mathbb{R}$ are also considered known, and where $y_E, y_E^{(i)}; i = 1, \ldots n-1$, and x_E are the estimated outputs, output derivatives, and the estimated states, over the estimation window $[a,b]$, respectively.

Remarks

• The estimates are exact on any estimation window in the following sense. In the absence of the measurement noise, the estimated output y_E is the exact system output in response to input u_K. Consequently, the estimated derivatives of the output and the system states are also exact and instantaneous.

• The algebraic estimator has a natural degree of robustness to measurement noise as y_E is obtained by the the action of an integral filter on the measured output signal y_M. Noise attenuation properties of the estimator are already good in the case of Gaussian additive measurement noise which is shown on many examples where the noise is "averaged out"; see also the results in [10].

• In practical applications the length of the estimation window may be adjusted to serve particular goals. To avoid accumulation of numerical errors during integration on long estimation horizons it is advisable to work with a short sliding window that is dragged forwards as new measurements arrive, much like it is done with receding horizon observers. This can be done seamlessly as the singularities are now removed from the kernels. Additionally, in the absence of model error, using partly overlapping estimation windows allows to assess robustness to noise by comparing the estimated trajectories on the overlapping intervals.

• In the present version of the estimator, and also in that of [9], the input signal (and its time derivatives) must be known a priori or have to be pre-filtered, if measured online. This is because the input u and its derivatives $u^{(k)}$ enter the equations for the estimates of output derivatives (49) in an un-filtered, direct way. It is important to note that equation (7) for the evaluation of output derivatives needs to hold only "almost everywhere". Consequently, lack of differentiability or even continuity of the system input on sets of measure zero will have no adverse effect on state estimation.

IV. EXAMPLE

The performance of the novel estimator is assessed on a simplified model of a DC motor, similar to that used in [9]. The model equations are:

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= \frac{-1}{t+1000} x_2(t) + \frac{1}{t+1000} u(t) \end{aligned}$$

with $y = x_1$ as measured output, x_1 as the angular position of the rotor, x_2 as the angular velocity of the rotor, and u as the control input voltage.

The controlled system invariance is:

$$y^{(2)}(t) + \frac{1}{t+1000} y^{(1)}(t) - \frac{1}{t+1000} u(t) = 0 \tag{50}$$

where the known input function is taken as: $u(t) = 120 \sin(t)$.

The state estimator parameters Γ and M in (49) are calculated as: $\Gamma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $M = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

The estimation horizon used here is $[0, 20s]$ and the estimator window is also chosen as $[0, 20s]$. The results show exact estimation of the system output and its derivative in the absence of measurement noise on the entire estimation horizon; see Figures 1, 2. Even if the output is perturbed with Gaussian noise with SNR (Signal-to-Noise Ratio) = 20 (see Figure 3), the estimation error is remarkably small and is uniformly bounded; see Figures 4, 5.

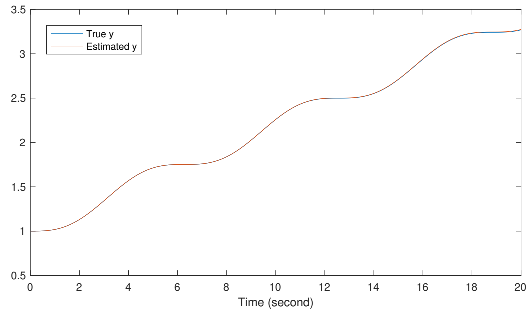

Fig. 1. Estimated output y for noiseless case

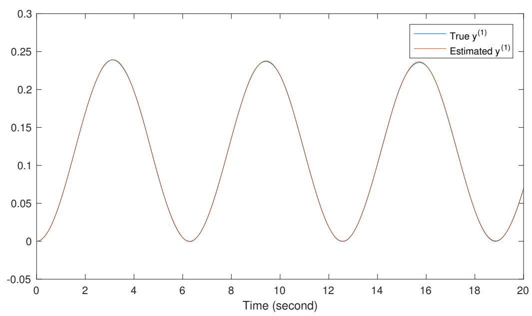

Fig. 2. Estimated derivative of y for noiseless case

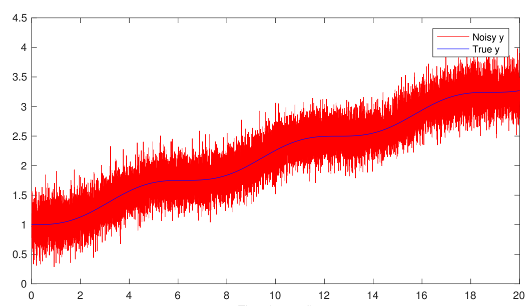

Fig. 3. Noisy y versus the true output y

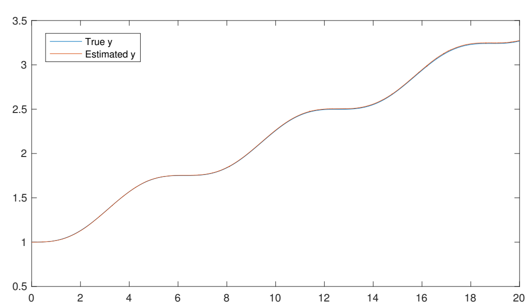

Fig. 4. Estimated output y for 20 SNR noise

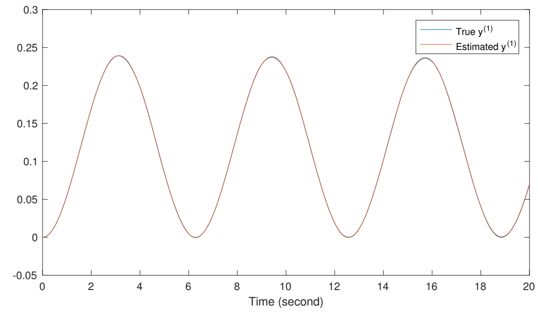

Fig. 5. Estimated derivative of y for 20 SNR noise

V. CONCLUSIONS

The integral system representation proposed here can be used in terms of local models in non-asymptotic state estimation in nonlinear systems. The kernel system representation has yet another advantage as stochastic processes also have kernel representations in reproducing kernel Hilbert spaces. When the stochastic process representing additive noise has known characteristics the two representations can be combined towards producing annihilators that can enhance robustness to measurement error. The algebraic estimation approach lends itself to further generalizations, notably it applies to MIMO systems. Details will be presented elsewhere.

REFERENCES

[1] M. Fliess, J. Lévine, P. Martin, and P. Rouchon, “Flatness and defect of non-linear systems: introductory theory and examples,” *International journal of control*, vol. 61, no. 6, pp. 1327–1361, 1995.

[2] M. Fliess, M. Mboup, H. Mounier, and H. Sira-Ramirez, “Questioning some paradigms of signal processing via concrete examples,” in *Algebraic Methods in Flatness, Signal Processing and State Estimation*. Editorial Lagares, 2003, pp. pp–1.

[3] D. P. Ghoshal, K. Gopalakrishnan, and H. Michalska, “Using invariance to extract signal from noise,” in *American Control Conference*, 2017.

[4] M. Fliess and H. Sira-Ramirez, “Control via state estimations of some nonlinear systems,” in *IFAC Symposium on Nonlinear Control Systtems (NOLCOS 2004)*, 2004.

[5] M. Fliess, C. Join, and H. Sira-Ramirez, “Non-linear estimation is easy,” *International Journal of Modelling, Identification and Control*, vol. 4, no. 1, pp. 12–27, 2008.

[6] H. Sira-Ramírez, C. G. Rodríguez, J. C. Romero, and A. L. Juárez, *Algebraic identification and estimation methods in feedback control systems*. John Wiley & Sons, 2014.

[7] M. Mboup, C. Join, and M. Fliess, “A revised look at numerical differentiation with an application to nonlinear feedback control,” in *The 15th Mediterrean Conference on Control and Automation-MED’2007*, 2007.

[8] ——, “Numerical differentiation with annihilators in noisy environment,” *Numerical algorithms*, vol. 50, no. 4, pp. 439–467, 2009.

[9] Y. Tian, T. Floquet, and W. Perruquetti, “Fast state estimation in linear time-varying systems: an algebraic approach,” in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*. IEEE, 2008, pp. 2539–2544.

[10] M. Fliess, “Analyse non standard du bruit,” *Comptes Rendus Mathematique*, vol. 342, no. 10, pp. 797–802, 2006.

[11] E. Kreindler and P. Sarachik, “On the concepts of controllability and observability of linear systems,” *IEEE Transactions on Automatic Control*, vol. 9, no. 2, pp. 129–136, 1964.

[12] L. M. Silverman and H. Meadows, “Controllability and observability in time-variable linear systems,” *SIAM Journal on Control*, vol. 5, no. 1, pp. 64–73, 1967.