2017 American Control Conference
Sheraton Seattle Hotel
May 24–26, 2017, Seattle, USA

Using Invariance to Extract Signal from Noise*

Debarshi Patanjali Ghoshal[1], Kumar Gopalakrishnan[1] and Hannah Michalska[2]

***Abstract*— It is shown how differential invariance can be used to extract an underlying signal from its noisy measurement towards constructing a non-asymptotic state estimator for linear systems. While the model of the system is assumed known, the noise can have arbitrary characteristics. The differential invariance is rendered by the Cayley-Hamilton theorem and the system is represented in terms of a output reproducing functional on a Hilbert subspace. High accuracy, full state estimation of the system is achieved over arbitrary time intervals by way of orthogonal projection onto the subspace that represents the system invariance. Although the results are presented here primarily with reference to SISO LTI systems they readily extend to LTV systems with multiple outputs.**

I. INTRODUCTION

The work presented here was inspired by Fliess [1]–[4], Sira-Ramirez [5], Perruquetti [6], Mboup [7], [8] et al., who first proposed an operational-calculus based, algebraic approach to high-order differentiation of analog signals. Real-time signal differentiation is of great importance if only for the reason that the knowledge of system output derivatives is necessary for algebraic state estimation in differentially flat nonlinear systems; [5]. For flat systems there exist system output functions $y = h(x); y \in \mathbb{R}^p$ such that the system state $x \in \mathbb{R}^n$ can be expressed as an algebraic mapping of the outputs and inputs $u \in \mathbb{R}^m$ of the system and their time derivatives, i.e. $x = F(y, \dot{y}, ..., y^{(k)}, u, \dot{u}, ..., u^{(k)})$, with k as the "order" of system flatness. The allure of the algebraic nonlinear estimators is attributed to the fact that the state estimates are direct and instantaneous (non-asymptotic) - a critical advantage in state estimation of hybrid systems where switching can occur at high rates; see Reger [9]. However, from a practical viewpoint, computing accurate high-order derivatives is considered unrealistic in noisy environments. Hence, asymptotic, approximate approaches such as the extended or unscented Kalman filters and particle filters, are still considered as the "reliable" options. The latter are not without their own weakness such as no guaranteed rate of convergence, strict assumptions about the system noise, or necessity of skillful tuning. With the development of new methods for high accuracy differentiation such preference will be less obvious.

Attempts to improve the methods of algebraic high order differentiation thus seem highly desirable and motivate the current developments presented here. First, it is noted that the original approaches of Fliess and Sira-Ramirez [1]–[5] pertain to differentiation of general signals f whose "models" are selected as truncated Taylor series expansions, otherwise stated as $\frac{d^n}{dt^n} f(t) = 0$ with n as the order of series truncation. The approach has two caveats: (i) the computation of estimates requires division by the time variable t that imposes singularity at $t = 0$; (ii) the method is not well adapted to handle noise albeit providing a certain degree of robustness by virtue of the fact that differentiation is "substituted" by integration that features a degree of natural filtering. This general approach is further examined by Mboup [7], [8] who points out that the error in signal derivative estimation has two sources: the model error (finite truncation of the modeling series) and the error due to measurement noise. The latter component is then reduced by making use of the freedom in selection of the integrating derivative estimator that is represented as a convolution operator. It is also pointed out that the approach is essentially akin to polynomial fitting to noisy signals.

As is the case in most target tracking applications, the signals to be differentiated are almost always associated with some underlying dynamics of systems that generate them. Hence, some degree of knowledge about the dynamic behaviour of the generating system can be reasonably assumed. It is important to point out that such knowledge almost never extends to any estimate of the initial conditions of the system. For the purpose of the present introductory note, full behavioural knowledge of the system is assumed as is done in Peruquetti [6].

With this motivation, the contributions presented here are summarized as follows. The system state equations are first replaced by a behavioural model of the system in the form of an output reproducing property on some arbitrary time horizon $[t_a, t_b]$. The behavioural model is easily derived from the differential invariance by which the system is characterized and eliminates the need for initial conditions of the system. The latter are in fact replaced by the properties of the system trajectories on $[t_a, t_b]$. More accurately, the model proposed is of the form of a homogeneous Fredholm integral equation of the second kind with a Hilbert-Schmidt kernel. The model kernel induces corresponding integral kernels for high order differentiation and does not exhibit any singularities. The complications associated with division by t at $t = 0$, present in all above cited works [1]–[8] are thereby eliminated. An equivalent mathematical interpretation of the behavioural model as a closed subspace of the kernel induced RKHS (reproducing kernel Hilbert space; see [10]) permits extraction of the signal and its time derivatives that uphold the validity of the system invariance from output

*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

978-1-5090-5994-2/$31.00 ©2017 AACC

measurement subject to arbitrary noise.

By virtue of the validity of the generalized Cayley-Hamilton's principle, the approach extends naturally to LTV MIMO systems; details will be presented elsewhere due to limited length of this article.

II. A BEHAVIOURAL MODEL OF AN LTI SYSTEM

A. Underlying Model Assumptions

In this introductory note, it is assumed that the signal to be extracted and differentiated is generated by a known and observable linear SISO system in state space form

$$\dot{x} = Ax + bu \ ; \quad y = c^T x \ ; \ x \in \mathbb{R}^n \tag{1}$$

For simplicity of exposition, but without loss of generality, it is also assumed that the influence of the exogenous input u on the system output y is factored out prior to state estimation and that the system characteristic equation is

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0 \tag{2}$$

Implied is an instantaneous recovery of the state $x(t)$ from the knowledge of the output $y(t)$ and its derivatives $y^{(i)}(t); i = 1, ..., n-1$ at t.

B. A Differential Invariant

In essence, a differential invariant under the action of a Lie group is a composite function

$$\mathcal{I}(t, y(t), y^{(1)}(t), \cdots, y^{(n)}(t)) \ ; \quad t \geq 0 \tag{3}$$

of time, and a system variable y with its first n time derivatives whose value is constant under the action of the group (action of the flow of the system (1)). Clearly, one such function is delivered by the left hand side of the characteristic equation (2) due to the validity of the Cayley-Hamilton's principle which ascertains that the system matrix satisfies

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 = 0 \tag{4}$$

Left-multiplication by c^T and right-multiplication by $x(t)$ yields

$$\begin{aligned} c^T A^n x(t) + a_{n-1} c^T A^{n-1} x(t) + \cdots \\ + a_1 c^T A x(t) + a_0 c^T x(t) = 0 \end{aligned} \tag{5}$$

that immediately translates into

$$y^{(n)}(t) + a_{n-1} y^{(n-1)}(t) + \cdots + a_1 y^{(1)}(t) + a_0 y(t) = 0 \tag{6}$$

so that $\mathcal{I}$ can be identified with the left hand side of (6).

C. Derivation of the Behavioural Model

Given an arbitrary time interval $[t_a, t_b] := [a, b]$ we shall strive to develop a surrogate model for (1) that instantly retrieves initial conditions of the system from its trajectories on $[a, b]$. Such an embedding may seem a highly inefficient way for reconstruction of the system state, if not that it also delivers efficient integral formulae for computation of all the derivatives of the system output. To this end we note the following result

Theorem 1: There exist Hilbert-Schmidt kernels K_{DS}, K^i_{DS}, $i = 1, \cdots n-1$, such that the output function y of (1) is reproduced on the interval $[a, b]$ in accordance with the action of the evaluation functional

$$y(t) = \int_a^b K_{DS}(t, \tau) y(\tau) \ \mathrm{d}\tau \ ; \ \forall t \in [a, b] \tag{7}$$

and the derivatives of the output $y^{(1)}, \cdots y^{(n-1)}$ can be computed *recursively* by way of output integration, so that for $i = 1, \cdots n-1$ and for all $t \in [a, b]$:

$$y^{(i)}(t) = \sum_{k=0}^{i-1} b_k(t) y^{(k)}(t) + \int_a^b K^i_{DS}(t, \tau) y(\tau) \ \mathrm{d}\tau \tag{8}$$

where $y^{(0)} \equiv y$ and $b_k(\cdot)$ are rational functions of t. 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 maybe somewhat cumbersome their derivation is surprisingly straightforward. The formal proof of the theorem is hence replaced by an example explicit, but abbreviated, derivation conducted with reference to a 3-dimensional system which is complex enough to clarify the procedure pointing to a complete proof by induction. To this end, let $n = 3$, and consider two equations obtained from (2) by pre-multiplication by the respective factors $(\xi - a)$ and $(b - \zeta)$:

$$(\xi - a)^3 y^{(3)} + a_2 (\xi - a)^3 y^{(2)} + a_1 (\xi - a)^3 y^{(1)} + a_0 (\xi - a)^3 y = 0 \tag{9}$$

$$(b - \zeta)^3 y^{(3)} + a_2 (b - \zeta)^3 y^{(2)} + a_1 (b - \zeta)^3 y^{(1)} + a_0 (b - \zeta)^3 y = 0 \tag{10}$$

Each of the above are then integrated three times, term by term, on the respective intervals $[a, a+\tau]$ and on $[b-\sigma, b]$ while assuming that τ and σ are related by $a + \tau = b - \sigma$. Integration by parts is used whenever it allows to lower the degree of the derivatives appearing under the integrals and the result is then simplified algebraically before proceeding to the next integration. To illustrate this process, integrating the first term of (9) yields

$$\begin{aligned} \int_a^{a+\tau} (\xi - a)^3 y^{(3)}(\xi) \, \mathrm{d}\xi &= (\xi - a)^3 y^{(2)}(\xi) \, \Big|_a^{a+\tau} - \int_a^{a+\tau} 3(\xi - a)^2 y^{(2)}(\xi) \, \mathrm{d}\xi \\ &= \tau^3 y^{(2)}(a+\tau) - \Big[3(\xi - a)^2 y^{(1)}(\xi) \, \Big|_a^{a+\tau} - \int_a^{a+\tau} 6(\xi - a) y^{(1)}(\xi) \mathrm{d}\xi \Big] \\ &= \tau^3 y^{(2)}(a+\tau) - 3\tau^2 y^{(1)}(a+\tau) + 6\tau y(a+\tau) - \int_a^{a+\tau} 6 y(\xi) \, \mathrm{d}\xi \end{aligned} \tag{11}$$

Integrating again gives

$$\int_a^{a+\tau}\int_a^{\xi'}(\xi-a)^3y^{(3)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'$$
$$=\tau^3y^{(1)}(a+\tau)-6\tau^2y(a+\tau)$$
$$+\int_a^{a+\tau}18(\xi'-a)y(\xi')\,\mathrm{d}\xi'-\int_a^{a+\tau}\int_a^{\xi'}6y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi' \qquad (12)$$

Integrating the third time yields

$$\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}(\xi-a)^3y^{(3)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''$$
$$=\tau^3y(a+\tau)-\int_a^{a+\tau}9(\xi''-a)^2y(\xi'')\,\mathrm{d}\xi'' \qquad (13)$$
$$+\int_a^{a+\tau}\int_a^{\xi''}18(\xi'-a)y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi''-\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}6y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''$$

To abbreviate derivations, we give the end results after triple consecutive integration of the remaining terms in (9)

$$\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}a_2(\xi-a)^3y^{(2)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''$$
$$=\int_a^{a+\tau}a_2(\xi''-a)^3y(\xi'')\,\mathrm{d}\xi''$$
$$-\int_a^{a+\tau}\int_a^{\xi''}6a_2(\xi'-a)^2y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi''$$
$$+\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}6a_2(\xi-a)y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \qquad (14)$$

The term containing the second derivative yields

$$\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}a_1(\xi-a)^3y^{(1)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''$$
$$=\int_a^{a+\tau}\int_a^{\xi''}a_1(\xi'-a)^3y(\xi')\,\mathrm{d}\xi'$$
$$-\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}3a_1(\xi-a)^2y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \qquad (15)$$

Finally, the last term is

$$\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}a_0(\xi-a)^3y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \qquad (16)$$

Collecting the terms in (11) - (16) yields

$$-\tau^3y(a+\tau)=\int_a^{a+\tau}\Big[-9(\xi''-a)^2$$
$$+a_2(\xi''-a)^3\Big]y(\xi'')\,\mathrm{d}\xi''$$
$$+\int_a^{a+\tau}\int_a^{\xi''}\Big[18(\xi'-a)-6a_2(\xi'-a)^2$$
$$+a_1(\xi'-a)^3\Big]y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi''$$
$$+\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}\Big[-6+6a_2(\xi-a)$$
$$-3a_1(\xi-a)^2+a_0(\xi-a)^3\Big]y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \qquad (17)$$

Applying the Cauchy formula for repeated integration while letting $a+\tau=t$ gives

$$(t-a)^3y(t)\triangleq\int_a^t K_F(t,\tau)y(\tau)\,\mathrm{d}\tau \qquad (18)$$

with $K_F(t,\tau)$ defined as

$$K_F(t,\tau)\triangleq\Big[9(\tau-a)^2-a_2(\tau-a)^3\Big] \qquad (19)$$
$$+(t-\tau)\Big[-18(\tau-a)+6a_2(\tau-a)^2-a_1(\tau-a)^3\Big]$$
$$+\frac{(t-\tau)^2}{2}\Big[6-6a_2(\tau-a)+3a_1(\tau-a)^2-a_0(\tau-a)^3\Big]$$

Equation (10) is processed similarly, with the end result obtained after collecting terms and applying the Cauchy formula while letting $b-\sigma=t$

$$(b-t)^3y(t)\triangleq\int_t^b K_B(t,\tau)y(\tau)\mathrm{d}\tau \qquad (20)$$

with $K_B(t,\tau)$ given by

$$K_B(t,\tau)\triangleq\Big[9(b-\tau)^2+a_2(b-\tau)^3\Big] \qquad (21)$$
$$+(t-\tau)\Big[18(b-\tau)+6a_2(b-\tau)^2+a_1(b-\tau)^3\Big]$$
$$+\frac{(t-\tau)^2}{2}\Big[6+6a_2(b-\tau)+3a_1(b-\tau)^2+a_0(b-\tau)^3\Big]$$

Adding (19) and (21) side by side while dividing both sides by $[(t-a)^3+(b-t)^3]$ yields

$$y(t)=\int_a^b K_{DS}(t,\tau)y(\tau)\,\mathrm{d}\tau \qquad (22)$$

with

$$K_{DS}(t,\tau) \triangleq \frac{1}{[(t-a)^3+(b-t)^3]}\begin{cases} K_F(t,\tau) & : \tau \le t \\ K_B(t,\tau) & : \tau > t \end{cases} \tag{23}$$

This delivers the formula for the "double-sided" kernel that effectively combines the operations of forward integration on $[a,t]$ and backward integration on $[b,t]$.

The recursive expressions for the derivatives of the output (8) can be derived by proceeding similarly as when deriving the K_{DS}. To obtain the expression for $y^{(1)}$ the equations (9) & (10) need to be integrated two times :

$$\begin{aligned}&(t-a)^3 y^{(1)}(t) = 6(t-a)^2 y(t) - a_2(t-a)^3 y(t) \qquad (24)\\ &+\int_a^t \big[-18(\tau-a)+6a_2(\tau-a)^2 - a_1(\tau-a)^3\big]y(\tau)\,\mathrm{d}\tau\\ &+\int_a^t (t-\tau)\big[6-6a_2(\tau-a)+3a_1(\tau-a)^2\\ &-a_0(\tau-a)^3\big]y(\tau)\,\mathrm{d}\tau\end{aligned}$$

and

$$\begin{aligned}&(b-t)^3 y^{(1)}(t) = -6(b-t)^2 y(t) - a_2(b-t)^3 y(t) \qquad (25)\\ &+\int_t^b \big[18(b-\tau)+6a_2(b-\tau)^2 + a_1(b-\tau)^3\big]y(\tau)\,\mathrm{d}\tau\\ &+\int_t^b (t-\tau)\big[6+6a_2(b-\tau)+3a_1(b-\tau)^2\\ &+a_0(b-\tau)^3\big]y(\tau)\,\mathrm{d}\tau\end{aligned}$$

The final expression for $y^{(1)}$ is obtained by adding the results of (24) and (25) while dividing by $[(t-a)^3+(b-t)^3]$.

To obtain a formula for $y^{(2)}$, equations (9) & (10) need to be integrated only once :

$$\begin{aligned}&(t-a)^3 y^{(2)}(t) = 3(t-a)^2 y^{(1)}(t) - a_2(t-a)^3 y^{(1)}(t) \qquad (26)\\ &-6(t-a)y(t)+3a_2(t-a)^2 y(t) - a_1(t-a)^3 y(t)\\ &+\int_a^t \big[6-6a_2(\tau-a)+3a_1(\tau-a)^2 - a_0(\tau-a)^3\big]y(\tau)\,\mathrm{d}\tau\end{aligned}$$

and

$$\begin{aligned}&(b-t)^3 y^{(2)}(t) = -3(b-t)^2 y^{(1)}(t) - a_2(b-t)^3 y^{(1)}(t) \qquad (27)\\ &-6(b-t)y(t)-3a_2(b-t)^2 y(t) - a_1(b-t)^3 y(t)\\ &+\int_t^b \big[6+6a_2(b-\tau)+3a_1(b-\tau)^2 + a_0(b-\tau)^3\big]y(\tau)\,\mathrm{d}\tau\end{aligned}$$

The expression for $y^{(2)}$ is obtained by adding (26) and (27) while dividing by the factor $[(t-a)^3+(b-t)^3]$.

Remark 1: It follows from the construction that the resulting kernel K_{DS} does not exhibit any singularities at the extremities of the integration interval $[a,b]$ because $\min_t[(t-a)^3+(b-t)^3] = 0.25(b-a)^3 > 0$. An important consequence of the double sided integration is that the estimates of the derivatives of the output are equally accurate over the entire estimation interval $[a,b]$.

It is also worth noting that the reproducing kernel induces a reproducing kernel Hilbert space (RKHS) uniquely corresponding to the symmetric, positive-type kernel function $K(t_1,t_2) \triangleq \langle K_{DS}(t_1,\cdot)|K_{DS}(t_2,\cdot)\rangle_2$ for all $t_1,t_2 \in [a,b]$ where $\langle\cdot|\cdot\rangle_2$ denotes the scalar product in $L^2[a,b]$. The RKHS, here denoted by $\mathcal{H}_K$, is then simply defined as the image of the space $L^2[a,b]$ under the integral transform defined by the double-sided kernel K_{DS} of (7) with a K-weighted norm as defined in [10]. The reproducing equality (7) has then yet another useful interpretation - that of a linear subspace of $\mathcal{H}_K$:

$$\mathcal{S}_I \triangleq \{y \in \mathcal{H}_K \mid y \text{ satisfies (7) }\} \tag{28}$$

More importantly, by construction of the double sided output reproducing kernel, it also follows that the behavioural model of Theorem 1 is equivalent to the differential model as described by the invariance equation (6). Precisely, we can state this as follows:

Corollary 1: An output function $y : [a,b] \to \mathbb{R}$ satisfies the invariance equation (6) on the interval $[a,b]$ if an only if it is reproduced by the evaluation functional in (7).

The proof of this fact is redundant as the multiple iterated integration in the derivation of the reproducing kernel can be reversed by multiple differentiation to retrieve the original invariance equation. Clearly, at this point, the initial conditions of the original system play no role as the behaviour of the system is fully characterized in terms of its trajectories over $[a,b]$. The filtering problem for the system output and its derivatives of any order then amounts to the application of Theorem 1 in a cascade manner:

$$y_E(t) = \int_a^b K_{DS}(t,\tau)y_M(\tau)\ \mathrm{d}\tau\ ;\ \forall t \in [a,b] \tag{29}$$

where the estimates for the derivatives of the output for $i = 1,\cdots n-1$ for $t \in [a,b]$ are calculated from:

$$y_E^{(i)}(t) = \sum_{k=0}^{i-1} b_k(t) y_E^{(k)}(t) + \int_a^b K_{DS}^i(t,\tau)y_E(\tau)\ \mathrm{d}\tau \tag{30}$$

in which y_M and y_E are the measured and estimated signals, respectively, and where $y_E^{(0)} = y_E$. As noted before the above estimation approach is already quite robust with respect to measurement noise due to the natural smoothing properties of the integral functionals that act like filters. In the absence of singularities the estimates are uniformly precise over the estimation interval whose length can be chosen arbitrarily.

III. OUTPUT ESTIMATION BY PROJECTION

Clearly, the system output can, in alternative to (29), be smoothed/reconstructed from a noisy measurement by direct orthogonal projection onto the subspace $\mathcal{S}_I$ which unambiguously embodies the conservation of the system invariant (6). It is further transparent that the subspace $\mathcal{S}_I$ is closed as it is finite dimensional and spanned by the fundamental set of solutions for the invariance equation (6). Such fundamental set, $\bar{y}_i(\cdot),\ i = 1,\cdots,n$, is readily obtained by direct integration of (6) for all unit vector initial conditions

$\bar{y}_i(a) = e_i; e_i \triangleq [0, \cdots, 0, 1, 0, \cdots 0] \in \mathbb{R}^n; i = 1, \cdots, n$. Due to linear independence of the latter

$$\mathcal{S}_I = \text{span}\{\bar{y}_i(\cdot),\ i = 1, \cdots, n\} \tag{31}$$

For a measured noisy signal $y_M(\cdot) \in L^2[a, b]$ its projection onto $\mathcal{S}_I$ is obtained as

$$y_E(\cdot) \triangleq \text{arg min}\{\| y_M - y \|_2^2 \ \mid y \in \mathcal{S}_I\} \tag{32}$$

The procedure is standard; putting

$$y_E = \sum_{i=1}^{n} \hat{c}_i \bar{y}_i \tag{33}$$

optimality in (32) is achieved if and only if

$$\langle y_M | \bar{y}_j \rangle_2 = \sum_{i=1}^{n} \hat{c}_i \langle \bar{y}_i | \bar{y}_i \rangle_2 \ j = 1, \cdots, n \tag{34}$$

which re-writes in a matrix form as

$$v = G(\bar{y})\hat{c}\ ; \quad G(\bar{y}) \triangleq \text{mat}\{\langle \bar{y}_i | \bar{y}_j \rangle_2\}_{i,j=1}^n \tag{35}$$
$$v \triangleq \text{vec}\{\langle y_M | \bar{y}_i \rangle_2\}_{i=1}^n\ ;\ \hat{c} \triangleq \text{vec}\{\hat{c}_i\}_{i=1}^n$$

where the Gram matrix G is invertible by linear independence of the spanning set of $\mathcal{S}_I$, thus yielding the optimal coefficient vector :

$$\hat{c} = G^{-1}(\bar{y})v \tag{36}$$

The estimated output y_E is then given by (33) and can be used in the recursive formulae (30) to deliver the corresponding estimates of the output derivatives.

It is very important to note the clear advantage of using formula (29). *In the absence of measurement noise, the estimate y_E of* (29) *is exact as it is identical to that obtained by projection* (33)*, yet it does not require the knowledge of the fundamental set (eigenvectors of the characteristic equation).* Many other advantages of the kernel system representation in an RKHS will be discussed elsewhere.

IV. EXAMPLE

A 3rd order LTI system was considered as follows:

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -10 & 0 \end{bmatrix} x\ ; y = x_1\ ;\ x(0) = [1, 1, 0] \tag{37}$$

The output y was subjected to different types of noise prior to estimation of $y, \dot{y}, \ddot{y}$ while using the kernels of Theorem 1. Romberg integration was used in calculation of the estimates. The results were compared with the estimates produced by a standard Kalman filter employing system model (37) with incorrect (unknown) initial conditions $x(0) = [-0.5, -0.5, -0.5]$.

A. Case I: Quantization with additional noise of 30 SNR

The simulated system output was first quantized using Lloyd's algorithm. Gaussian noise of 30 SNR was added to the quantized signal and the resulting noisy "measurement" was then used to compute the estimates $y, \dot{y}, \ddot{y}$. The noisy measurement and the estimation results are shown in Figures 1-4, respectively.

B. Case II: Purple noise

The inverse frequency power of purple noise is equal to -2. This type of noise was added to the output signal y before estimation. The noisy measured output is shown in Fig. 5 and the estimation results are then compared in Figures 6-8. The initial conditions employed by the Kalman filter were again different from the true ones used to generate the system output.

V. CONCLUSIONS

The constructed kernel filter has some obvious advantages over recursive asymptotic observers: it delivers estimates in arbitrarily limited time; the estimates are exact in a noise free case; it has a degree of robustness to noise as integral smoothing is involved; and it does not need initiation by estimates of the initial conditions of the system.

The specific filter kernel constructed here has yet another advantage over its competitors. It delivers uniformly accurate estimates over the entire estimation interval as the potential singularities at the extremities of the estimation interval are removed by a combination of forward and backward integration. Finally, the kernel model of the system captures invariance as a geometrical object - an interpretation that will prove useful in extending the ideas introduced here not only to LTV but also to nonlinear systems.

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

[4] 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.

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

[6] 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.

[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] M. Mboup, C. Join, and M. Fliess, "Numerical differentiation with annihilators in noisy environment," *Numerical algorithms*, vol. 50, no. 4, pp. 439–467, 2009.

[9] J. Reger, H. Sira-Ramirez, and M. Fliess, "On non-asymptotic observation of nonlinear systems," in *Proceedings of the 44th IEEE Conference on Decision and Control*. IEEE, 2005, pp. 4219–4224.

[10] S. Saitoh, *Theory of reproducing kernels and its applications*. Longman, 1988, vol. 189.

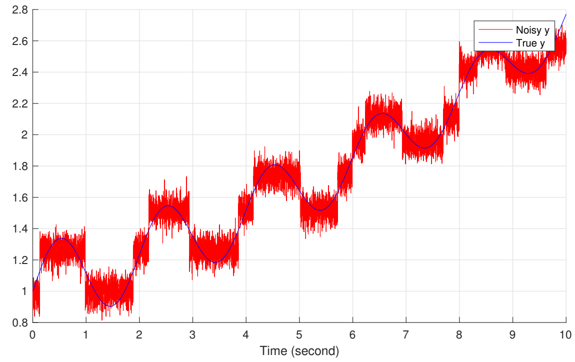

Fig. 1. Case I: Quantized, noisy y versus the true output y.

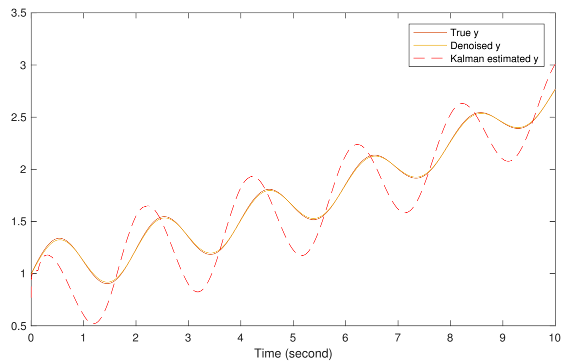

Fig. 2. Case I: Estimated output y by kernel vs. Kalman filter.

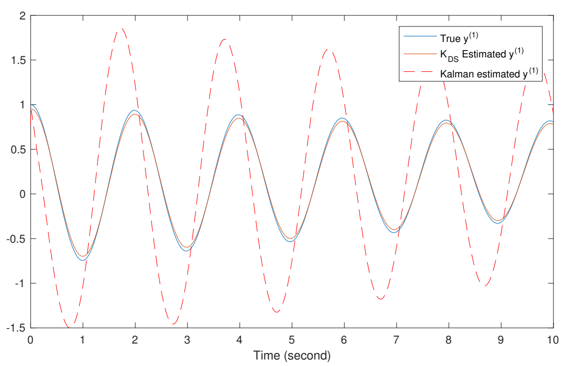

Fig. 3. Case I: Estimated first derivative of y by kernel vs. Kalman filter.

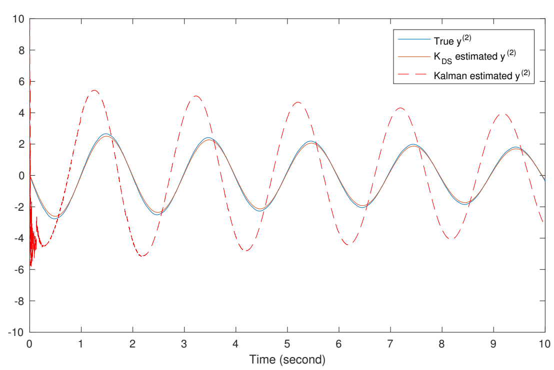

Fig. 4. Case I: Estimated second derivative of y by kernel vs. Kalman filter.

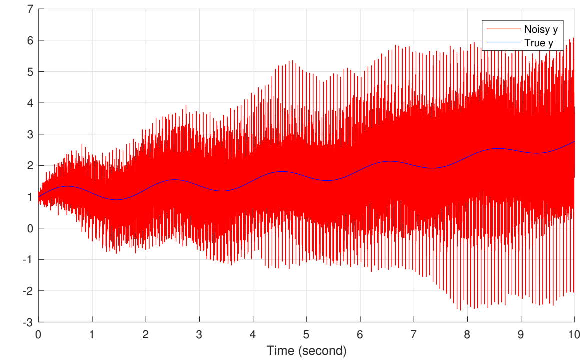

Fig. 5. Case II: Noisy (purple noise) y versus the true output y.

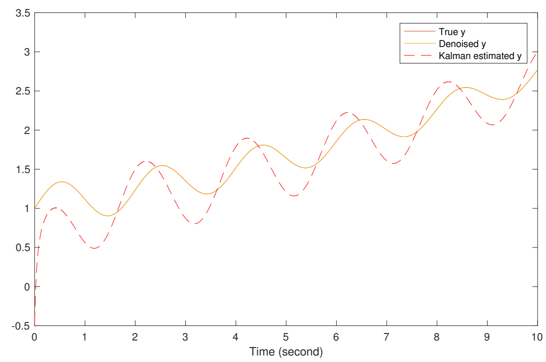

Fig. 6. Case II: Estimated output y by kernel vs. Kalman filter.

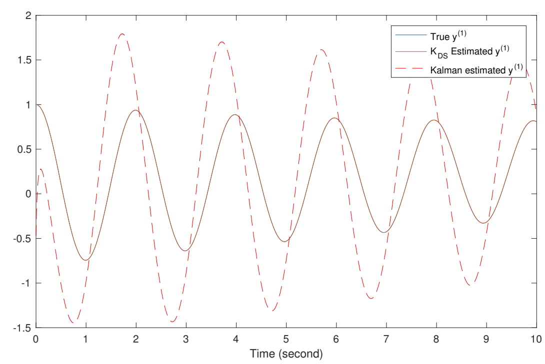

Fig. 7. Case II: Estimated first derivative of y by kernel vs. Kalman filter.

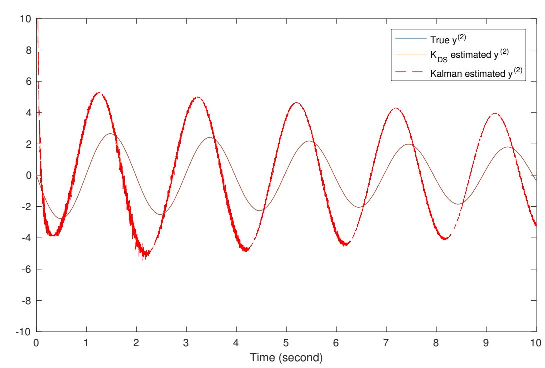

Fig. 8. Case II: Estimated second derivative of y by kernel vs. Kalman filter.