

Non-Asymptotic Numerical Differentiation: a Kernel-Based Approach

Peng Li^a, Gilberto Pin^b, Giuseppe Fedele^c and Thomas Parisini^{a,d,e*}

^aDepartment of Electrical and Electronic Engineering, Imperial College London, UK;

^bElectrolux Professional S.p.A., Italy; ^cDepartment of Informatics, Modeling, Electronics and Systems Engineering, University of Calabria, Italy ; ^dDepartment of Engineering and Architecture, University of Trieste, Italy; ^e KIOS Research and Innovation Center of Excellence, Cyprus

ARTICLE HISTORY

Compiled May 15, 2018

ABSTRACT

The derivative estimation problem is addressed in this paper by using Volterra integral operators which allow to obtain the estimates of the time-derivatives with fast convergence rate. A deadbeat state observer is used to provide the estimates of the derivatives with a given fixed-time convergence. The estimation bias caused by modeling error is characterized herein as well as the ISS property of the estimation error with respect to the measurement perturbation. A number of numerical examples are carried out to show the effectiveness of the proposed differentiator also including comparisons with some existing methods.

KEYWORDS

Linear integral operators, Numerical differentiation, Non-Asymptotic identification, State estimation, Fredholm-Volterra integral equations

1. Introduction

The design of robust numerical differentiators is a long-standing problem due to its significant role in applied mathematics and control engineering, for instance in parameter estimation (Chai, Lin, & Fu, 2013), controller implementation (Fliess & Sira-Ramirez, 2004), (Levant, 2003) and so forth.

The methods available in the literature tend to form four mainstreams:

- *least-squares polynomial fitting or interpolation* (see Duncan, Mandl, and Pasik-Duncan (1996) and Ibrir and Diop (2004)) which is a very classical method regularizing the parameters through optimization showing efficiency in off-line applications;
- *digital filtering in frequency domain*, which inherits, adopting a signal-processing perspective, the use of the finite impulse response filter, see Chen and Lee (1995), and infinite impulse response digital filter as proposed in Rader and Jackson (2006) to capture the derivatives of the objective signal;

This work has been partially supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 739551 (KIOS CoE).

* Corresponding author. Email: t.parisini@gmail.com

- *dynamic-based derivative estimation*, which, as the name implies, making use of the signal dynamics to track the unmeasurable derivatives dynamically, taking the advantages of continuous-time differential algebra and operational calculus, etc.

Finally, the *Sliding Mode* approach, based on discontinuous injection, is a powerful tool that is widely resorted to. Examples can be found in (Bartolini, Pisano, & Usai, 2000) and (Levant, 2003). In the SM framework, high-gain injection is inevitable which will degrade the performance while facing the disturbance. Moreover, SM methods suffer from either the chattering problem (low-order SM such as Bartolini et al. (2000)) or high complexity in implementation (higher-order SM).

In the recent literature, the deadbeat estimation has gained increasing attention due to its finite-time or fixed-time convergence properties. The algebraic methodology has been applied to the differentiation problem in Fliess and Sira-Ramirez (2004) making use of symbolic calculus and consecutive integrations to provide a good estimate of the first derivatives. As a further improvement, Mboup, Join, and Fliess (2007) provides a new efficient approach of derivative estimation, known as the Jacobian differentiator, which solves the problem by means of parameter estimation. Further relevant contributions are Liu, Gibaru, and Perruquetti (2011a) and Liu, Gibaru, and Perruquetti (2011b). In Reger and Jouffroy (2009), the authors recast the derivative estimation problem in an state observation fashion, making use of a deadbeat algebraic state estimation method to provide the derivative estimates.

Inspired by the perspective of Reger and Jouffroy (2009), in this paper, we are extending the application of the kernel-based observer to solve the numerical differentiation problem. The kernel-based approach was first proposed to face the parameter estimation problem (Pin, Assalone, Lovera, & Parisini, 2016), where taking advantages of Volterra integral operational algebra and the non-asymptotic property of the kernels, the effects of unknown initial conditions which cause transient process are removed thus achieving instantaneous convergence. Remarkably, the kernel-based deadbeat estimation methodology has been recently extended to the state observer design (Pin, Lovera, Assalone, & Parisini, 2013). Implemented with a series of Bivariate Feedthrough Non-asymptotic Kernels (BF-NK), the kernel-based observer is able to reconstruct the states of an observable LTI system in arbitrarily short time. Utilizing the kernel-based observer, the main idea of the present paper consists in modeling the signal as a truncated Taylor series, for which a trivial generator system (model) exists, that can be used for the design of a deadbeat state observer.

This paper is organized as follows: The relationship between the derivative estimation and the state observation is formulated with corresponding state-space realization in Section 2. In Section 3, the algebra of the Volterra integral operator and the non-asymptotic kernel are briefly reviewed. The ad-hoc algorithm of the kernel-based derivative estimator and its implementation are presented in Section 4. Then, in Section 5, the truncation error and the robustness of the proposed differentiator are characterized and extensive numerical examples are carried out in Section 6. The last section is devoted to conclusions.

2. From state observer to time-derivative estimation

Let $y(t)$ be a real-valued analytic signal whose derivatives are expected to be estimated from its observation on a finite-time interval. At $t = 0$, the signal $y(t)$ admits the Taylor

series expansion

$$y(t) = \sum_{i=0}^{\infty} \frac{y^{(i)}(0)}{i!} t^i, \quad (1)$$

where $y^{(i)}(t)$ denotes its i -th derivative. Consider the truncated approximation of the Taylor expansion

$$y_N(t) = y_N(0) + y_N^{(1)}(0)t + \frac{y_N^{(2)}(0)}{2}t^2 + \dots + \frac{y_N^{(N-1)}(0)}{(N-1)!}t^{N-1}, \quad (2)$$

which verifies the differential equation $y^{(N)}(t) = 0$. The truncated signal $y_N(t)$ can be assumed generated by the following LTI system

$$\begin{cases} \dot{x}(t) = Ax(t), \\ y_N(t) = cx(t), \quad t \in \mathbb{R}_{\geq 0}, \end{cases} \quad (3)$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{N \times N}, \quad c = [1 \ 0 \ \dots \ 0] \in \mathbb{R}^N, \quad (4)$$

with the state vector $x(t) = [y_N(t), y_N^{(1)}(t), \dots, y_N^{(i)}(t), \dots, y_N^{(N-1)}(t)]^\top$.

Remarkably, the system (3) is completely observable. In turn, replacing $y_N(t)$ with $y(t)$, state observation of system (3) provides the possibility to estimate the derivatives up to $(N-1)$ -th order of the signal $y(t)$ from its measurements.

3. Volterra linear integral operator's theory and non-asymptotic kernel

The kernel-based observer, as the key tool of the present derivative estimator, is based on the Volterra linear integral operator. For the readers' convenience, the basic concepts and useful algebra are briefly recalled here. For detailed operational theory, readers may refer to Burton (2005).

Let $r(t) \in \mathbb{R}$, $\forall t \geq 0$ be an i -th order differentiable signal. Given a Hilbert-Schmidt (\mathcal{HS}) kernel function $K(\cdot, \cdot)$, in two variables, its i -th order derivative with respect to the second argument is denoted as $K^{(i)}(\cdot, \cdot)$, $i \in \mathbb{Z}_{\geq 0}$. The Volterra integral operator induced by the kernel function is

$$[V_K r](t) \triangleq \int_0^t K(t, \tau) r(\tau) d\tau, \quad t \in \mathbb{R}_{\geq 0}. \quad (5)$$

Remarkably the transformed signal $[V_K r](t)$, for $t \geq 0$, can be obtained as the output

of the following LTV dynamic system:

$$\begin{cases} \zeta^{(1)}(t) &= K(t, t)r(t) + \int_0^t \left(\frac{\partial}{\partial t} K(t, \tau) \right) r(\tau) d\tau \\ [V_K r](t) &= \zeta(t) \end{cases} \quad (6)$$

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

Lemma 3.1. *Pin et al. (2016)* For a given $i \in \mathbb{Z}_{\geq 0}$, consider a signal $r(\cdot) \in \mathcal{L}^2(\mathbb{R})$ that admits i -th derivative and a \mathcal{HS} kernel function $K(\cdot, \cdot)$, admitting the i -th derivative with respect to the second argument. Then, it holds that:

$$\begin{aligned} [V_K r^{(i)}](t) &= \sum_{j=0}^{i-1} (-1)^{i-j-1} r^{(j)}(t) K^{(i-j-1)}(t, t) + \sum_{j=0}^{i-1} (-1)^{i-j} r^{(j)}(0) K^{(i-j-1)}(t, 0) \\ &\quad + (-1)^i [V_{K^{(i)}} r](t) \end{aligned} \quad (7)$$

□

Non-asymptotic properties of the Volterra operator can be achieved by designing a class of kernel functions according to the following definition.

Definition 3.2. If a \mathcal{HS} kernel $K(\cdot, \cdot)$ which is at least $(i-1)$ -th order differentiable with respect to the second argument, verifies the condition

$$K^{(j)}(t, 0) = 0, \forall j \in \{0, 1, \dots, i-1\} \quad (8)$$

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

Definition 3.3. (BIBO-stability of Volterra operator (Pin et al., 2016)) A Volterra operator V_K is said to be BIBO-stable if

$$\sup_{t \in \mathbb{R}_{>0}} \left\{ \int_0^t |K(t, \tau)| d\tau \right\} > \infty.$$

Remarkably, the non-asymptotic property suppresses the effect of unknown initial conditions, getting the differentiator rid of the transient to achieve a fast convergence. Compared to the mono-variate modulating functions adopted by modulating function-based observers (e.g. Pin, Chen, and Parisini (2015)), the proposed bivariate non-asymptotic kernels add to the non-asymptoticity also the BIBO stability property, which permits to avoid the periodic resetting that is mandatory for the modulating-function methods. As a result, the Volterra operators induced by the said kernels can be implemented as stable dynamical systems.

4. Fast-convergent derivative estimator

The proposed derivative estimation consists in estimating the internal states of a linear dynamic system

$$\begin{cases} \dot{\hat{x}}(t) = A\hat{x}(t), \\ y(t) = c\hat{x}(t), \quad t \in \mathbb{R}_{\geq 0}, \end{cases} \quad (9)$$

with A and c defined in (4).

Induced by an N -th or higher order non-asymptotic kernel function $K_h(t, \tau)$, recalling (7) and considering (8), the Volterra operator will map the output signal $y(t)$ and its derivatives into the transformation space as

$$\left[V_{K_h} y^{(i)} \right](t) = \sum_{j=0}^{i-1} (-1)^{i-j-1} y^{(j)}(t) K_h^{(i-j-1)}(t, t) + (-1)^i [V_{K_h^{(i)}} y](t), \quad i \in \{1, \dots, N-1\}. \quad (10)$$

As a result, the effect of the unknown initial conditions $y^{(j)}(0), \forall j \in \{0, \dots, i-1\}$ are annihilated.

Consider the $i = 1$ case, the above equation writes

$$[V_{K_h^{(1)}} y](t) = y(t) K_h(t, t) - [V_{K_h} y^{(1)}](t). \quad (11)$$

Similarly, the transformation also holds for $y^{(N-1)}(t)$ and takes the form

$$[V_{K_h^{(1)}} y^{(N-1)}](t) = y^{(N-1)}(t) K_h(t, t) - [V_{K_h} y^{(N)}](t). \quad (12)$$

Recall the fact that $y^{(N)}(t) = 0, \forall t \in \mathbb{R}_{>0}$ and consider the definition in (5), (12) reduces to

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

which, according to (10) is equivalent to

$$\begin{aligned} (-1)^{N-1} [V_{K_h^{(N)}} y](t) &= \sum_{j=0}^{N-1} (-1)^{N-j-1} y^{(j)}(t) K_h^{(N-j-1)}(t, t) \\ &= \sum_{j=0}^{N-1} (-1)^{N-j-1} K_h^{(N-j-1)}(t, t) \hat{x}_j(t), \end{aligned} \quad (13)$$

where $\hat{x}_j(t)$ denotes j th element of $\hat{x}(t)$.

Defining $\gamma_{h,j}(t) \triangleq (-1)^{N-j-1} K_h^{(N-j-1)}(t, t)$, (13) can be written in a compact form

$$(-1)^{N-1} [V_{K_h^{(N)}} y](t) = \sum_{j=0}^{N-1} \gamma_{h,j}(t) \hat{x}_j(t) = \gamma_h^\top \hat{x}(t), \quad (14)$$

with $\gamma_h = [\gamma_{h,N-1}(t), \gamma_{h,N-2}(t), \dots, \gamma_{h,0}(t)]^\top$.

We augment the above linear constraint by stacking N equations in the form of (13) induced by an array of Bivariate Feedthrough Non-asymptotic Kernel (BF-NK defined in Pin et al. (2013)) having the form

$$K_h(t, \tau) = e^{-\omega_h(t-\tau)} (1 - e^{-\bar{\omega}\tau})^N, h \in \{0, \dots, N-1\}, \quad (15)$$

sharing the same $\bar{\omega} \in \mathbb{R}_{>0}$ but with different $\omega_h \in \mathbb{R}_{>0}$, $h \in \{0, \dots, N-1\}$. Indeed, the condition (8) up to the N -th order is met by the factor $(1 - e^{-\bar{\omega}t})^N$. Therefore, the derivative estimation problem translates into solving for $\hat{x}(t)$ the following algebraic system

$$\nu(t) = \Gamma(t)\hat{x}(t), \quad (16)$$

where

$$\nu(t) = \left[(-1)^{N-1} \left[V_{K_0^{(N)}} y \right] (t), (-1)^{N-1} \left[V_{K_1^{(N)}} y \right] (t), \dots, (-1)^{N-1} \left[V_{K_{N-1}^{(N)}} y \right] (t), \right]^\top,$$

and

$$\Gamma(t) = [\gamma_0(t), \gamma_1(t), \dots, \gamma_{N-1}(t)]^\top.$$

Note that in (16), except the unknown state vector $\hat{x}(t)$, all the other terms consist of known functions of the kernels derivatives $K_h^{(j)}(t, t)$, $j \in \{0, \dots, N-1\}$. Moreover, the transformed signal $\xi(t) \triangleq \left[\left[V_{K_0^{(N)}} y \right] (t), \left[V_{K_1^{(N)}} y \right] (t), \dots, \left[V_{K_{N-1}^{(N)}} y \right] (t) \right]^\top = (-1)^{N-1} \nu(t)$, thanks to the shape of the kernel and (6), can be calculated by a LTV dynamic system

$$\begin{cases} \xi^{(1)}(t) &= G\xi(t) + E(t)y(t), \\ \xi(0) &= 0, \end{cases} \quad (17)$$

where

$$\begin{aligned} G &= \text{diag}(-\omega_h, h \in \{0, \dots, N-1\}) \in \mathbb{R}^{(N-1) \times (N-1)}, \\ E(t) &= \left[K_0^{(N)}(t, t), K_1^{(N)}(t, t), \dots, K_{N-1}^{(N)}(t, t) \right]^\top. \end{aligned}$$

Remark 1. The transformation system (17) is internally stable thanks to the chosen shape of the kernel (15) where there is no need for periodically re-initialization as in Fliess and Sira-Ramirez (2004) nor moving the time horizon as proposed in Reger and Jouffroy (2009).

Lemma 4.1. (*Invertibility*) *Given the kernels in the form of (15) tuned by a common $\bar{\omega}$ and different ω_h , $h \in \{0, \dots, N-1\}$, the matrix $\Gamma(t)$ will be invertible $\forall t \in \mathbb{R}_{>0}$.*

Proof. Recall (15), the BF-NK can be rearranged as

$$K_h(t, \tau) = e^{-\omega_h t} \sum_{q=0}^N (-1)^q \binom{N}{q} e^{(\omega_h - q\bar{\omega})\tau}. \quad (18)$$

Its i -th derivative with respect to the second argument takes the form

$$K_h^{(i)}(t, \tau) = e^{-\omega_h t} \sum_{q=0}^N (-1)^q \binom{N}{q} (\omega_h - q\bar{\omega})^i e^{(\omega_h - q\bar{\omega})\tau}. \quad (19)$$

Therefore, the vector $\gamma_h(t)^\top$ becomes

$$\begin{bmatrix} (-1)^{N-1} \binom{N}{0} (\omega_h - q\bar{\omega})^{N-1} & (-1)^N \binom{N}{1} (\omega_h - q\bar{\omega})^{N-1} & \dots & \binom{N}{N} (\omega_h - q\bar{\omega})^{N-1} \\ (-1)^{N-2} \binom{N}{0} (\omega_h - q\bar{\omega})^{N-2} & (-1)^N \binom{N}{1} (\omega_h - q\bar{\omega})^{N-2} & \dots & -\binom{N}{N} (\omega_h - q\bar{\omega})^{N-2} \\ \vdots & \vdots & & \vdots \\ \binom{N}{0} & -\binom{N}{1} & \dots & (-1)^N \binom{N}{N} \end{bmatrix} \begin{bmatrix} 1 \\ e^{-\bar{\omega}t} \\ \vdots \\ e^{-N\bar{\omega}t} \end{bmatrix}. \quad (20)$$

Moreover, each element of $\gamma_h(t)$ is persistently exciting, i.e. $\exists \epsilon_h \in \mathbb{R}_{>0}$ and $T_h \in \mathbb{R}_{>0}$ such that

$$\int_{t-T_h}^t \gamma_h(\tau) \gamma_h(\tau)^\top d\tau \geq \epsilon_h I, \quad \forall t \geq 0.$$

By taking different values of $\omega_h \in \mathbb{R}_{>0}$, referring to (20), $\Gamma(t)$ has a nonsingular determinant for any $t > 0$. In other words, the matrix $\Gamma(t)$ is invertible $\forall t \in \mathbb{R}_{>0}$. \square

Finally, with the persistently invertible $\Gamma(t)$, the estimated derivatives can be obtained by

$$\hat{x}(t) = \begin{bmatrix} \hat{y}(t) \\ \hat{y}^{(1)}(t) \\ \vdots \\ \hat{y}^{(N-1)}(t) \end{bmatrix} = \Gamma(t)^{-1} \nu(t), \quad t \in \mathbb{R}_{>0}. \quad (21)$$

5. Derivative estimation error analysis

5.1. Analysis of the Estimation error due to Taylor-expansion truncation

Compared to the simplified N -dimensional linear system (9) used for the design of the observer (considered as the generator for the truncated Taylor expansion of the observed signal), the full-model (exact) extension can be modeled by augmenting (9) with a further unknown input, that represents the overall contribution of higher-order derivatives.

$$\begin{cases} \dot{z}(t) &= Az(t) + bu_N(t) \\ y(t) &= cz(t), \quad \forall t \in \mathbb{R}_{>0}, \end{cases} \quad (22)$$

where $b = [0, \dots, 0, 1]^\top \in \mathbb{R}^N$ and $u_N(t) \triangleq R_N^{(N)}(t) \in \mathbb{R}$ with $R_N(t)$ denoting the truncation error of the Taylor series, i.e. $R_N(t) = y(t) - y_N(t)$. In (22), the state vector $z(t) = [y(t), y^{(1)}(t), \dots, y^{(N-1)}(t)]$ is made up of the true derivatives of the signal $y(t)$.

By applying the Volterra operators to the full-order dynamical system (22), we get an extended version of (12), that accounts for the additional input

$$\begin{aligned} [V_{K_h^{(1)}} y^{(N-1)}](t) &= y^{(N-1)}(t)K_h(t, t) - [V_{K_h} y^{(N)}](t) \\ &= y^{(N-1)}(t)K_h(t, t) - [V_{K_h} u_N](t) \end{aligned} \quad (23)$$

Therefore, after some tedious algebra, the linear constraint for the full order system can be expressed in the operators domain as:

$$(-1)^{N-1}[V_{K_h^{(N)}} y](t) + [V_{K_h} u_N](t) = \sum_{j=0}^{N-1} (-1)^{N-j-1} K_h^{(N-j-1)}(t, t) z_j(t). \quad (24)$$

The augmentation by N BF-NKs gives

$$\nu(t) + e_u(t) = \Gamma(t)z(t), \quad (25)$$

where

$$e_u(t) \triangleq \left[[V_{K_0} u_N](t), [V_{K_1} u_N](t), \dots, [V_{K_{N-1}} u_N](t) \right]^\top$$

is an unknown input vector due to the Taylor truncation.

In turn, the derivative estimation-error vector can be expressed as a function of e_u

$$e_a(t) \triangleq z(t) - x(t) = \Gamma(t)^{-1} e_u(t). \quad (26)$$

In the view of the structure of $\Gamma(t)$, its inverse writes

$$\Gamma(t)^{-1} = \frac{1}{\det(A)} \text{adj}(\Gamma(t)), \quad (27)$$

where $\text{adj}(\Gamma(t))$ denotes the adjoint matrix of $\Gamma(t)$. It is worth noting that in $\text{adj}(\Gamma(t))$, the elements on the j -th row are functions of derivatives of the kernel having order higher than that on the $(j-1)$ -th row, for all $j \in \{2, \dots, N\}$. For instance, let us consider the $N=3$ case ¹

$$\text{adj}(\Gamma) = \begin{bmatrix} K_1 K_2^{(1)} - K_1^{(1)} K_2 & -K_0 K_2^{(1)} + K_0^{(1)} K_2 & K_0 K_1^{(1)} - K_0^{(1)} K_1 \\ K_1 K_2^{(2)} - K_1^{(2)} K_2 & -K_0 K_2^{(2)} + K_0^{(2)} K_2 & K_0 K_1^{(2)} - K_0^{(2)} K_1 \\ K_1^{(1)} K_2^{(2)} - K_1^{(2)} K_2^{(1)} & -K_0^{(1)} K_2^{(2)} + K_0^{(2)} K_2^{(1)} & K_0^{(1)} K_1^{(2)} - K_0^{(2)} K_1^{(1)} \end{bmatrix}, \quad (28)$$

By choosing ω_h as large as to verify the inequality

$$\omega_h - q\bar{\omega} > 1, \quad \forall q \in \{0, \dots, N\}, \quad \forall h \in \{1, \dots, N-1\},$$

the absolute value of elements on the same column in $\Gamma(t)^{-1}$ increases row by row. As a result, recall (26), the estimation error will become larger for higher-order derivatives. Some instances will be given in the simulation section in the present paper.

¹For the sake of clarity, we eliminate the time-dependence of the time-varying matrix $\Gamma(t)$, the kernels $K(t, t)$ and their derivatives $K_h^{(i)}(t, t)$, $\forall t \in \{1, 2\}$, $\forall h \in \{1, 2\}$.

5.2. Perturbed scenario

Assume now that $y(t)$ is affected by a bounded additive measurement perturbation

$$y_n(t) = y(t) + d_y(t),$$

where $|d_y(t)| \leq \bar{d}_y$.

Due to the linearity of the Volterra operator, defining

$$e_y(t) \triangleq (-1)^{N-1} [[V_{K_0^{(N)}} d_y](t), [V_{K_1^{(N)}} d_y](t), \dots, [V_{K_{N-1}^{(N)}} d_y](t)]^\top,$$

the image vector of the perturbed measurements $y_n(t)$ can be expressed as

$$\xi_n(t) = \xi(t) + e_y(t), \quad (29)$$

where $e_y(t)$ admits the dynamics

$$\begin{cases} e_y^{(1)}(t) &= G e_y(t) + E(t) d_y(t), \\ e_y(0) &= 0, \end{cases} \quad (30)$$

with previously defined G and $E(t)$, for $h \in \{0, \dots, N-1\}$. Consequently, in this case, the estimate is given by

$$\hat{x}_n = \Gamma(t)^{-1} \left[\nu(t) + e_y(t) \right] (t). \quad (31)$$

Referring to (25), the estimation error in the perturbed scenario takes the form

$$e_n(t) \triangleq z(t) - \hat{x}_n(t) = \Gamma(t)^{-1} \left(e_u(t) - e_y(t) \right). \quad (32)$$

5.3. Boundedness of the estimation error

Considering the estimation based on the perturbed measurements $y_n(t)$, let us explore the boundedness of the derivative estimation error $e_n(t)$ in (32).

As far as the truncation error is concerned, the following result shows the boundedness of the derivative error.

Proposition 5.1. *Let $|y^{(N)}(t)| \leq M$, for some finite $M \in \mathbb{R}_{>0}$, then one can conclude that the estimation bias $e_a(t)$ (recall (26)) of the signal derivatives given by the estimator (21) has a global bound $\bar{e}_a \in \mathbb{R}_{>0}$, i.e.*

$$|e_a(t)| \leq \bar{e}_a, \quad \forall t \in \mathbb{R}_{>0}. \quad (33)$$

Proof. The truncation error R_N can be rewritten in the Lagrange form (Abramowitz & Stegun, 1964)

$$R_N(t) = \frac{y^{(N)}(c) t^N}{(N)!}, \quad (34)$$

for some $c \in (0, t)$. Moreover, if the N -th derivative $|y^{(N)}(t)| \leq M, \forall t \in \mathbb{R}_{>0}$, the Taylor remainder admits an upper bound that

$$R_N(t) \leq \bar{R}_N = \frac{Mt^N}{(N)!}. \quad (35)$$

In turn, the input of (22) verifies that $|u_N(t)| \leq M$. Recall the BIBO stable transformation system (6),

$$\begin{aligned} |[V_{K_h} u_N](t)| &= \left| \int_0^t e^{-\omega_h(t-\tau)} K(\tau, \tau) u_N(\tau) d\tau \right| \\ &\leq \int_0^t e^{-\omega_h(t-\tau)} \sup_{0 \leq \tau \leq t} |K(\tau, \tau)| M d\tau \\ &= \frac{1}{\omega_h} \sup_{0 \leq \tau \leq t} |K(\tau, \tau)| M \triangleq \zeta_h, \quad h \in \{0, 1, \dots, N-1\}. \end{aligned} \quad (36)$$

Due to the invertibility of $\Gamma(t)$ for all $t \in \mathbb{R}_{>0}$, the derivative estimation error caused by truncation can be bounded with

$$|e_a(t)| \leq \left| \sup_{0 \leq \tau \leq t} \Gamma(\tau)^{-1} \left| [\zeta_0, \zeta_1, \dots, \zeta_{N-1}]^\top \right| \right| \triangleq \bar{e}_a. \quad (37)$$

□

We are characterizing the error introduced by the measurement perturbation $e_y(t)$ in the following proposition.

Proposition 5.2. (*ISS property*) *Given the perturbed measurement $y_n(t)$ with a bounded measurement disturbance $d_y(t)$ such that $|d_y(t)| \leq \bar{d}_y, \forall t \in \mathbb{R}$, the estimation error $e_y(t)$ is ISS with respect to $d_y(t)$.*

Proof. Referring to the dynamics of $e_y(t)$ (30), being G Hurwitz, there exists a positive definite matrix P satisfying the identity $PG + G^\top P = -I$. Therefore, we consider the Lyapunov function candidate $V(e_y) = e_y^\top P e_y$, then there exist two positive constant α_1 and α_2 verifying²

$$\alpha_1 \|e_y\|^2 \leq V(e_y) \leq \alpha_2 \|e_y\|^2, \quad \forall e_y.$$

Differentiating the Lyapunov function $V(e_y)$ along the trajectory of the error dynamic, we get

$$\dot{V}(e_y) = \frac{\partial V}{\partial e_y} \dot{e}_y \leq -\|e_y\|^2 + 2\|E\| |d_y|. \quad (38)$$

For any $0 < \epsilon < 1$, define a function

$$\chi(|d_y|) \triangleq \left(\frac{2\|E\|}{1-\epsilon} |d_y| \right)^{\frac{1}{2}}.$$

²For notation simplicity, we eliminate the time-dependence of $e_y(t)$ and $E(t)$.

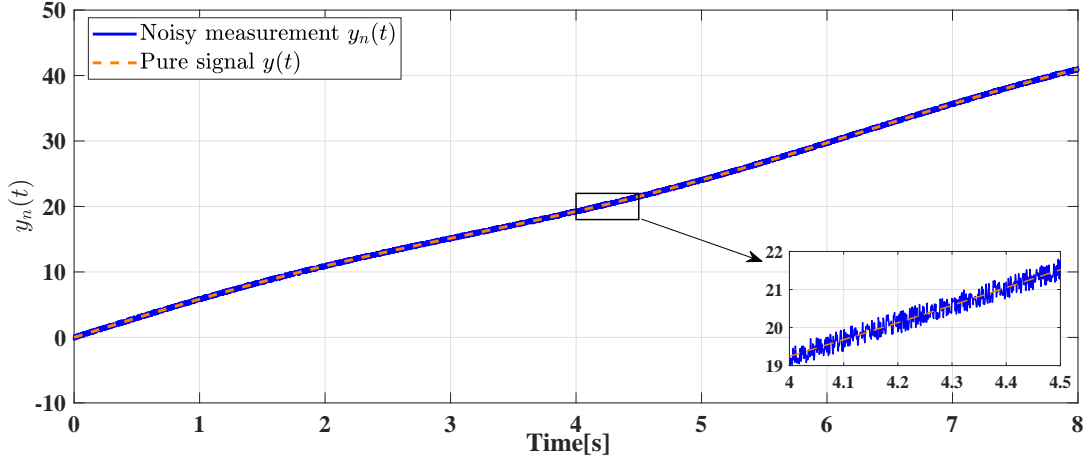


Figure 1. Pure signal $y(t)$ and the perturbed measurement $y_n(t)$.

It is obvious that

$$\dot{V} \leq -\epsilon \|e_y\|^2, \forall \|e_y\|^2 \geq \chi(|d_y|).$$

Thereby, it is straightforward to show that the estimation error caused by the measurement disturbance $e_y(t)$ is ISS with asymptotic gain

$$\gamma_{e_y}(|d_y|) = \alpha_1^{-1} \alpha_2 \chi(|d_y|), \quad (39)$$

indicating that for any arbitrary $\mu \in \mathbb{R}_{>0}$, $e_y(t)$ will enter a close ball of radius $\gamma_{e_y}(\bar{d}_y) + \mu$ in finite-time. \square

6. Numerical examples

In this section, we are going to verify the effectiveness of the proposed differentiator with extensive numerical experiments involving comparisons with some well-known differentiation methods.

6.1. First order differentiator ($N = 2$)

We consider, as first example, the derivative estimation of the drifting sinusoidal signal

$$y(t) = 5t + \sin(t),$$

in both ideal and perturbed scenario where the measurements are corrupted by a bounded random perturbation simulated as a uniformly distributed sequence ranging within $[-0.3, 0.3]$ (in this case the signal-to-noise ratio is $SNR \triangleq 10 \log_{10} \frac{\sum |y(k)|^2}{\sum |d_y(k)|^2} = 42.85$).

The perturbed measurement $y_n(t) = y(t) + d_y(t)$ and the pure signal $y(t)$ are shown in Fig. 1. The simulations are executed in Matlab/Simulink with a sampling interval $T_s = 10^{-3}s$.

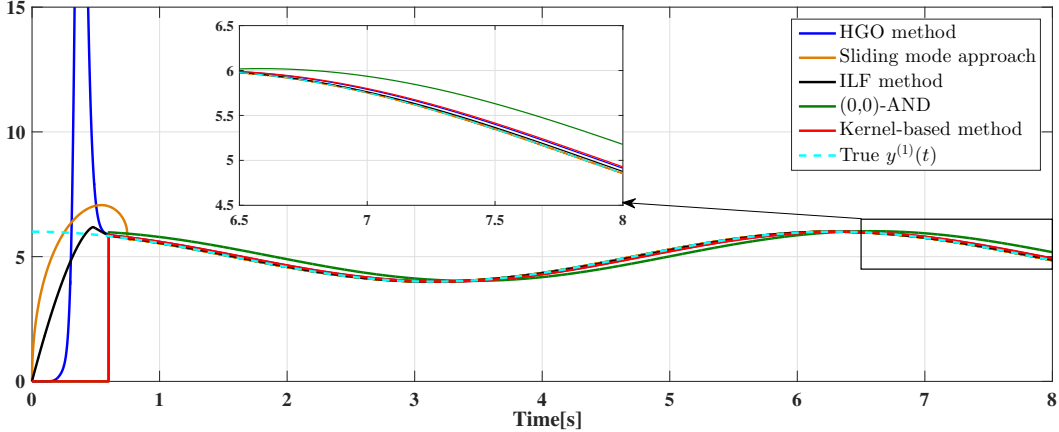


Figure 2. Estimates of first derivative $y^{(1)}(t)$ in perturbation-free scenario with $N = 2$.

We are estimating the first derivative of the signal from the observation $y(t)$ and its perturbed counterpart $y_n(t)$ using the estimator proposed in (21) with $N = 2$, tuned with parameters $\bar{\omega} = 2.5$ and $[\omega_0, \omega_1] = [1, 2] \times 20$. The estimation results of the proposed method are compared with some renowned approaches:

- (0, 0)-Algebraic Numerical Differentiator (AND) (Mboup, Join, & Fliess, 2009), with $M = 600$.
- Implicit Lyapunov Function (ILF) method (Polyakov, Efimov, & Perruquetti, 2014), with

$$P = \begin{bmatrix} 20.0608 & -2.6724 \\ -2.6724 & 0.3744 \end{bmatrix}, k = \begin{bmatrix} -14.0872 \\ -52.2407 \end{bmatrix}, \mu = 0.5;$$

- 2-sliding mode method (Levant, 1998) with $\alpha = 8$ and $\lambda = 6$;
- High-gain observer (HGO) method (Ibrir, 2004) with $\epsilon = 0.1$, $\lambda = 0.95$.

All the parameters are chosen by the criteria that all the estimators have similar convergence time (activation time for the deadbeat methods).

The estimates of the first order derivative of $y(t)$ provided by the four estimators are reported in Figure 2 based on the perturbation-free measurements. In Figure 2, one can conclude that the HGO method and the sliding mode approach have obvious overshoots during the transient time and reach the steady state at approximately $t = 0.6s$. Remarkably, after entering the steady state, the sliding mode method reconstructs the first derivative with the highest exactness among the methods. The ILF method, with comparatively tedious overshoot, provides steady estimates at around $t = 0.6s$. The deadbeat methods, (0, 0)-AND and the proposed kernel-based approach, activated at $t = 0.6s$, are able to track the variation of $y^{(1)}(t)$, albeit there is an obvious delay in the estimates from (0, 0)-AND, where the delay is not adjustable for fixed shape of annihilator and activation time (see Mboup et al. (2009)). However, the delay can be calculated. Generally speaking, all the differentiation approaches are able to provide acceptable estimates of the first derivative of $y(t)$ in steady state.

Under the disturbance of $d_y(t)$, the estimators perform the first derivative estimation as shown in Figure 3. The SM method has heightened fluctuation on its estimation results since the impact of the disturbance is amplified by the high-gain injection. The kernel-based and the HGO methods provide the estimates with bounded fluctuations

Table 1. Steady-state RMSE and $Var(E_n)$ of the differentiators in the perturbed scenario (SNR=42.85).

Method	HGO	Sliding mode	Kernel-based	ILF	(0,0)-AND
RMSE	37.084	164.21	28.689	9.3639	14.187
Variance	0.7541	5.8822	0.6444	0.5571	0.5324

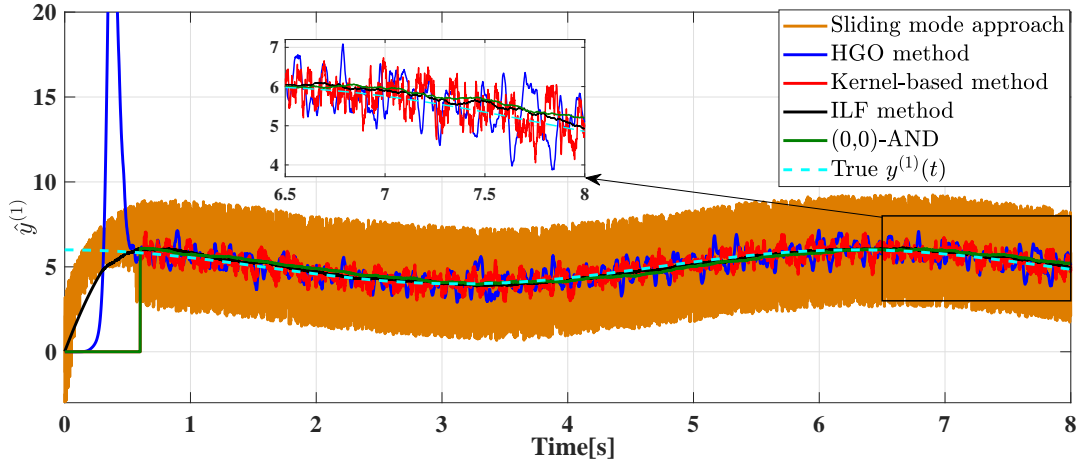


Figure 3. Estimates of first derivative $y^{(1)}(t)$ in perturbed scenario with $N = 2$.

within comparatively smaller ranges subsuming the true value, while the large overshoot in the transient phase of HGO method still exists. Notably, with the chosen parameters, the ILF method and the (0,0)-AND perform the estimations smoothly, although the (0,0)-AND has a computable delay.

In order to show the perturbation immunity of the simulated differentiators in a more intuitive way, we resort to the root mean square error (RMSE) defined as $RMSE \triangleq \sqrt{\frac{\sum_i^{M_k} (y^{(1)}(i) - \hat{y}^{(1)}(i))^2}{M_k}}$, where M_k denotes the number of data points considered and the variance of the estimation error $Var(e_n)$ in the steady state. The RMSE and the variance of estimates given by the differentiators on the time-interval $t \in [3, 8]$ are listed in Table 1, from which one can conclude that, with the present parameter settings, the ILF method and the algebraic method provide comparatively accurate and concentrated estimates in the presence of the perturbation on the measurements.

6.2. Second order differentiator ($N = 3$)

Increasing the dimension of the observer to 3, we are going to examine the performance of the proposed differentiator for simultaneously estimating the first and second derivatives of the signal. Moreover, in the previous simulation example, we observe that the kernel-based derivative estimator behaves mediocly in disturbance attenuation. Thanks to the theoretical formulation in Section 5, estimation error caused by the measurement perturbation is tunable with different parameter settings. Thereby, in this example, we also focus on the tuning rules of the parameters in the sense of the estimation bias and the robustness.

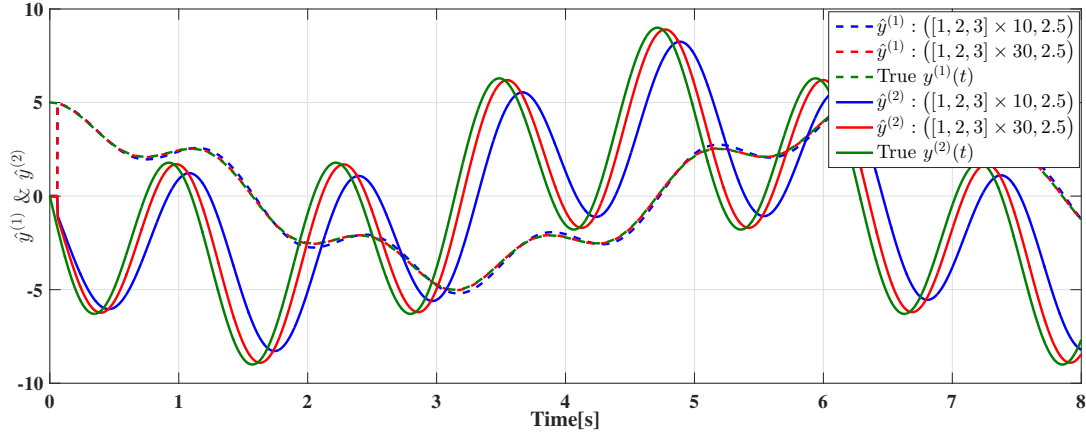


Figure 4. Estimates of first derivative $y^{(1)}(t)$ and the second derivative $y^{(2)}(t)$ in perturbation-free scenario with $N = 3$.

Consider a dual-sinusoidal signal

$$y(t) = 4 \sin(t) + 0.2 \sin(5t),$$

as depicted in Figure 5. The estimates of both the second order derivative $\hat{y}^{(2)}(t)$ and the first order derivative $\hat{y}^{(1)}(t)$ are depicted in Figure 4. The simulation is based on MATLAB/Simulink environment with the sampling interval $T_s = 10^{-3}s$.

In Figure 4, the derivative estimations of the proposed kernel-based differentiator are presented with different parameter configurations without consideration of the measurement perturbation. We choose two sets of parameters for $[\omega_0, \omega_1, \omega_2]$: $[1, 2, 3] \times 10$ and $[1, 2, 3] \times 50$ to show the effect of the parameter configuration.

In Figure 4, more evident in the estimates of $y^{(2)}(t)$, the differences between the estimates and the true values increase with the decreasing value of $[\omega_0, \omega_1, \omega_2]$ which is coincide with (26) where larger ω_h , $h \in \{0, 1, 2\}$, which can be also regarded as forgetting factors, will reduce the effect of the past data and thus provide more instantaneous and accurate estimates.

In the perturbed scenario, with a bounded perturbation $d_y(t)$ uniformly distributed in $[-0.2, 0.2]$ corresponding to an $SNR = 24.47$ (see Figure 5), the kernel-based differentiator gives the estimation results of $y^{(1)}(t)$ and $y^{(2)}(t)$ as shown in Figure 6. To observe the effect of ω_h on the estimation error clearly, in this example we consider more parameter settings: $[1, 2, 3] \times 10$, $[1, 2, 3] \times 20$, $[1, 2, 3] \times 30$ and $[1, 2, 3] \times 50$.

In this perturbed scenario, the estimator provides the estimates with fluctuations, where the range of the fluctuations are tunable with different parameters settings. As shown in Figure 6, the fluctuations are amplified as the value of $[\omega_0, \omega_1, \omega_3]$ increases, which verifies the robustness analysis in Subsection 5.2 due to ascending γ_{e_y} recall (39). According to (32) the value of $e_y(t)$ will be magnified with increasing value of ω_h , $h \in \{0, 1, 2\}$, thus enlarge $e_n(t)$.

In general, the parameters $\{\omega_h, h \in \{0, 1, 2\}\}$ should be selected considering the trading off between the bias and the disturbance immunity.

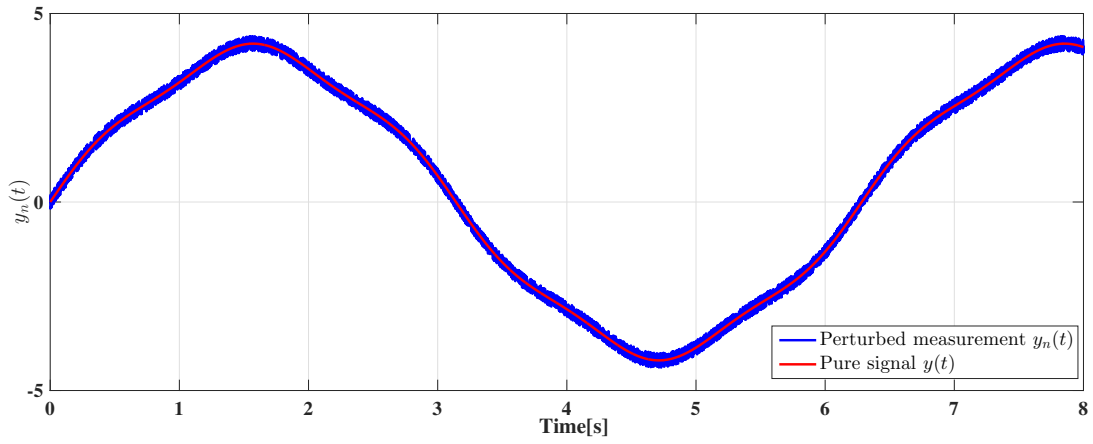


Figure 5. Pure signal $y(t)$ and the perturbed measurement $y_n(t)$.

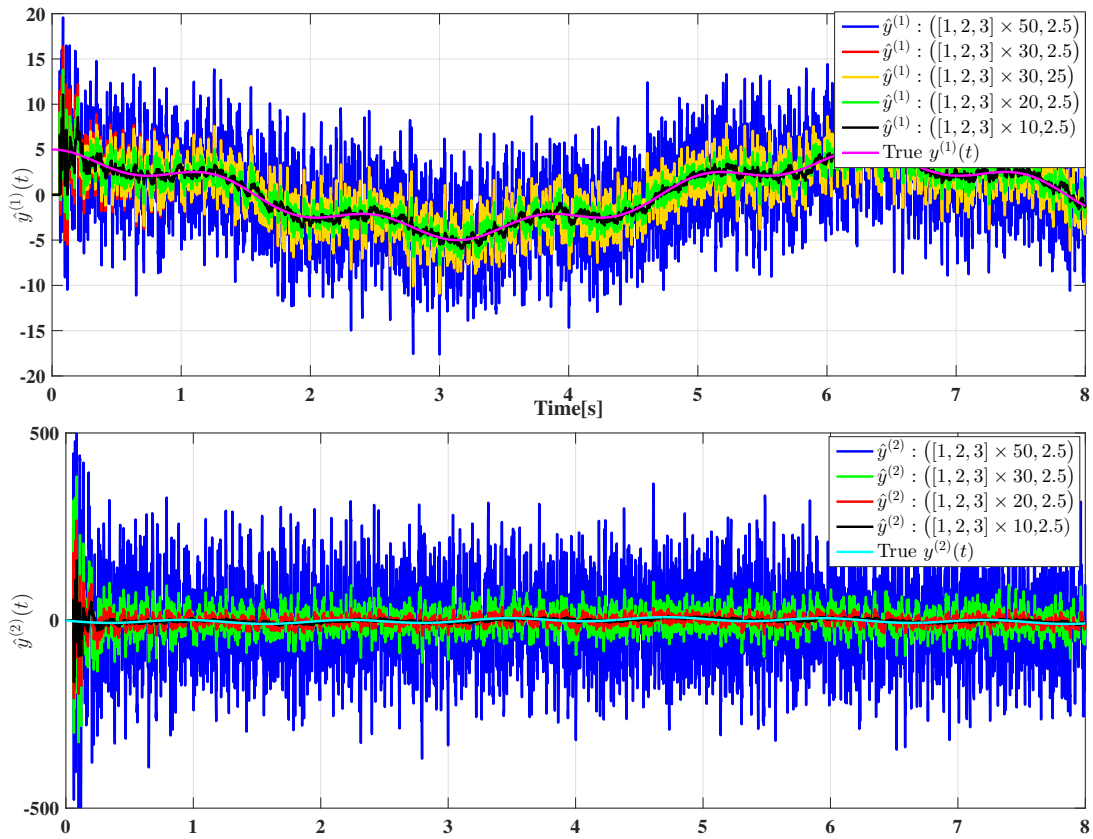


Figure 6. Estimates of first derivative $y^{(1)}(t)$, the second derivative $y^{(2)}(t)$ and error bound in perturbed scenario with $N = 3$ and different parameter settings

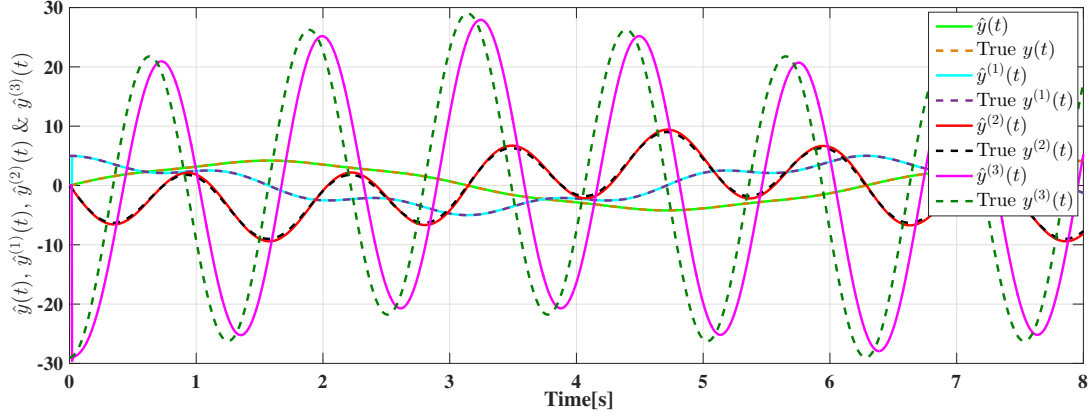


Figure 7. Estimates of derivatives of $y^{(i)}(t)$, $i \in \{0, 1, 2, 3\}$ in perturbation-free scenario with $N = 4$.

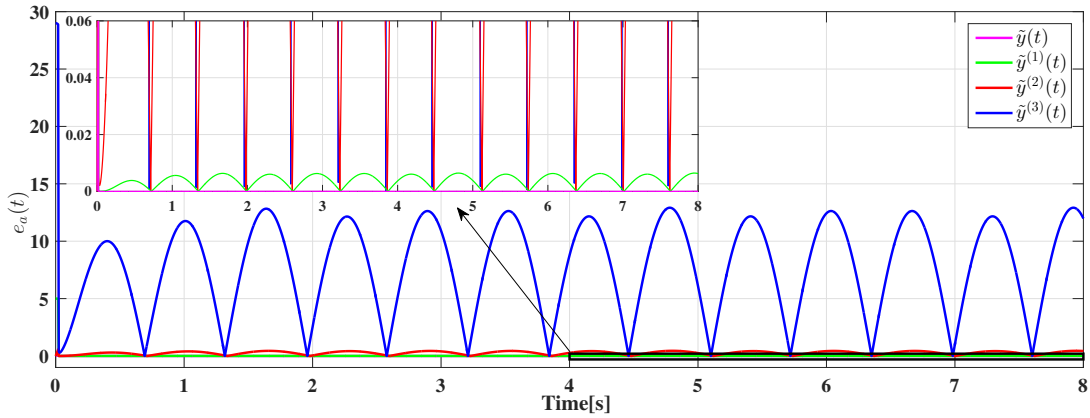


Figure 8. Derivative estimates error $e_{a,y^{(i)}}(t)$, $i \in \{0, 1, 2, 3\}$ in perturbation-free scenario with $N = 4$.

6.3. Third order differentiator ($N = 4$)

In this example, the same signal as in Subsection 6.2 is used to estimate up to the third derivative by using a fourth-order kernel-observer. From Figure 4, readers may notice that, with fixed parameter settings, the estimation bias depends on the different orders of derivatives. Thereby, this example will examine the value of estimation error with respect to the derivative levels, corresponding to Subsection 5.1. The parameters of the kernel-based differentiator are chosen as $[\omega_0, \omega_1, \omega_2] = [1, 2, 3] \times 20$, $\bar{\omega} = 2.5$ simulated in Matlab/Simulink with a sampling interval $T_s = 10^{-3}s$.

The perturbation-free derivative estimation results are reported in Figure 7 and corresponding estimation errors are shown in Figure 8. For further analysis, we consider the steady state value

$$\bar{\Gamma} \triangleq \lim_{t \rightarrow \infty} \Gamma(t) = \begin{bmatrix} -8000 & 400 & -20 & 1 \\ -64000 & 1600 & -40 & 1 \\ -216000 & 3600 & -60 & 1 \\ -512000 & 6400 & -80 & 1 \end{bmatrix}. \quad (40)$$

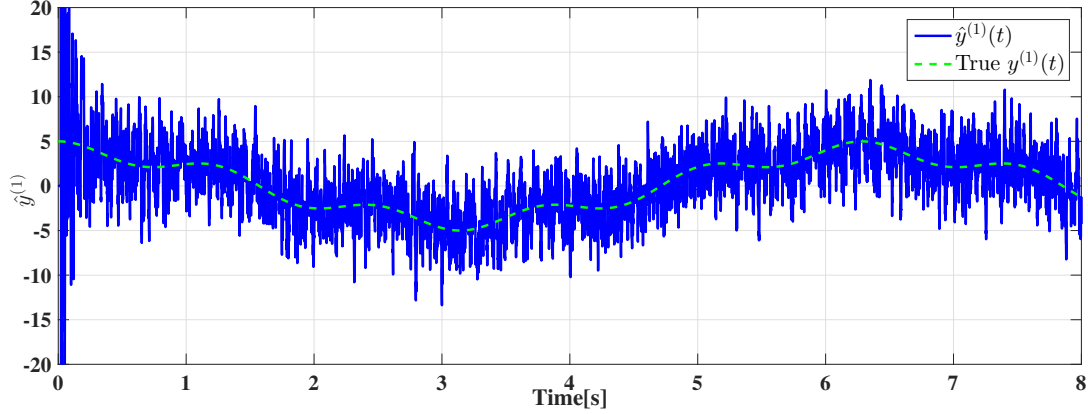


Figure 9. Estimates of derivatives of $y^{(1)}(t)$ in perturbed scenario with $N = 4$.

In turn, its inverse is

$$\bar{\Gamma}^{-1} = \begin{bmatrix} 0 & -0.0001 & 0.0001 & 0 \\ 0.0038 & -0.0100 & 0.0088 & -0.0025 \\ 0.2167 & -0.4750 & -0.3500 & -0.0917 \\ 4 & -6 & 4 & -1 \end{bmatrix}. \quad (41)$$

Right multiplying $\bar{\Gamma}^{-1}$ to the modeling error vector $e_u(t)$, the values of the estimation bias of the derivatives (elements of vector $e_a(t) = [e_{a,y}, e_{a,y^{(1)}}, e_{a,y^{(2)}}, e_{a,y^{(3)}}]^\top$) rank in the following order

$$e_{a,y^{(3)}} > e_{a,y^{(2)}} > e_{a,y^{(1)}} > e_{a,y}. \quad (42)$$

Similarly as in Subsection 6.2, in the perturbation case, the differentiator provides robust estimates with bounded fluctuations as reported in Figure 9 where, for simplicity and clearness, only the estimates of $y^{(1)}(t)$ are plotted. As we can see, the kernel-based numerical differentiator is able to give the derivative estimates with bounded error.

7. Concluding remarks

In this work, we have shown that the kernel-based observer methodology (Pin et al., 2013) can be applied to solve the numerical differentiation problem, obtaining fixed-time convergence of the estimated derivatives. Making use of the Volterra integral operators, induced by suitably shaped kernel functions, the derivatives of the signal can be tracked based on the measurements with a rapid convergence. Estimation bias and the robustness against the measurement perturbation are characterized herein. It can be proved that the estimation error is ISS with respect to the signal derivatives per se and the bounded perturbation.

Extensive numerical examples have been presented to examine the effectiveness of the proposed numerical differentiation method involving comparisons with existing renowned differentiation methods and echoing the theoretical analysis and derivation.

Further research efforts will concentrate on more comprehensive analysis on the robustness issue and improve the immunity of the proposed differentiator from the

aspects of parameter tuning and the shape of the kernel functions and potentially the numerical realization.

References

- Abramowitz, M., & Stegun, I. A. (1964). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables* (Vol. 55). Courier Corporation.
- Bartolini, G., Pisano, A., & Usai, E. (2000). First and second derivative estimation by sliding mode technique. *Journal of Signal Processing*, 4(2), 167–176.
- Burton, T. (2005). *Volterra integral and differential equations*. Elsevier.
- Chai, G., Lin, Z., & Fu, M. (2013). Consensus-based cooperative source localization of multi-agent systems. In *Control conference (ccc), 2013 32nd chinese* (pp. 6809–6814).
- Chen, C.-K., & Lee, J.-H. (1995). Design of high-order digital differentiators using $1/\text{sub } 1/\text{error}$ criteria. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 42(4), 287–291.
- Duncan, T. E., Mandl, P., & Pasik-Duncan, B. (1996). Numerical differentiation and parameter estimation in higher-order linear stochastic systems. *IEEE Transactions on Automatic control*, 41(4), 522–532.
- Fliess, M., & Sira-Ramirez, H. (2004). Control via state estimations of some nonlinear systems. In *Ifac symposium on nonlinear control systems (nolcos 2004)*.
- Ibrir, S. (2004). Linear time-derivative trackers. *Automatica*, 40(3), 397–405.
- Ibrir, S., & Diop, S. (2004). A numerical procedure for filtering and efficient high-order signal differentiation. *International Journal of Applied Mathematics and Computer Science*, 14(2), 201–208.
- Levant, A. (1998). Robust exact differentiation via sliding mode technique. *automatica*, 34(3), 379–384.
- Levant, A. (2003). Higher-order sliding modes, differentiation and output-feedback control. *International journal of Control*, 76(9-10), 924–941.
- Liu, D.-Y., Gibaru, O., & Perruquetti, W. (2011a). Differentiation by integration with jacobi polynomials. *Journal of Computational and Applied Mathematics*, 235(9), 3015–3032.
- Liu, D.-Y., Gibaru, O., & Perruquetti, W. (2011b). Error analysis of jacobi derivative estimators for noisy signals. *Numerical Algorithms*, 58(1), 53–83.
- Mboup, M., Join, C., & Fliess, M. (2007). A revised look at numerical differentiation with an application to nonlinear feedback control. In *Control & automation, 2007. med'07. mediterranean conference on* (pp. 1–6).
- Mboup, M., Join, C., & Fliess, M. (2009). Numerical differentiation with annihilators in noisy environment. *Numerical algorithms*, 50(4), 439–467.
- Pin, G., Assalone, A., Lovera, M., & Parisini, T. (2016, Feb). Non-asymptotic kernel-based parametric estimation of continuous-time linear systems. *IEEE Transactions on Automatic Control*, 61(2), 360-373.
- Pin, G., Chen, B., & Parisini, T. (2015). The modulation integral observer for linear continuous-time systems. In *Control conference (ecc), 2015 european* (pp. 2932–2939).
- Pin, G., Lovera, M., Assalone, A., & Parisini, T. (2013). Kernel-based non-asymptotic state estimation for linear continuous-time systems. In *2013 american control conference* (pp. 3123–3128).
- Polyakov, A., Efimov, D., & Perruquetti, W. (2014). Homogeneous differentiator design using implicit lyapunov function method. In *Control conference (ecc), 2014 european* (pp. 288–293).
- Rader, C. M., & Jackson, L. B. (2006). Approximating noncausal iir digital filters having arbitrary poles, including new hilbert transformer designs, via forward/backward block recursion. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 53(12), 2779–2787.

Reger, J., & Jouffroy, J. (2009). On algebraic time-derivative estimation and deadbeat state reconstruction. In *Decision and control, 2009 held jointly with the 2009 28th chinese control conference. cdc/ccc 2009. proceedings of the 48th ieee conference on* (pp. 1740–1745).