2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

A Comparison of Methods for Higher-order Numerical Differentiation

Kim D. Listmann, Zheng Zhao

Abstract— **This article compares three different methods for the online computation of higher-order derivatives from a measurement signal. In general such measurements are noise corrupted and the application of finite difference schemes is inappropriate. Thus, all methods are compared w.r.t. their noise suppression capabilities for different noise types and levels, their computational complexity to compute a derivative and their tuning effort for proper commissioning. Finally, a recommendation is provided which of the differentiators is best used when.**

I. Introduction

In many industrial automation processes the computation of derivatives of measurement signals is common. Be it for simply replacing a sensor or for estimating certain system parameters. The applications range from feedback control systems over data acquisition to simulation and modeling in dynamic systems [7]. Recently there is an increasing interest in data-driven feedback control [9] and model-free control [6] with the intent to define broadly applicable control laws that still perform well and are easy to maintain from an engineering point of view. Then, differentiation algorithms are a major subcomponent that decide about the efficiency and accuracy of the control scheme.

Classically, differentiation is done using finite difference schemes in combination with low-pass filters accounting for the measurement noise. This introduces inadvertent delays and seems unsuitable in particular for the computation of higher-order derivatives due to the sensitivity to measurement noise. However, there exist more sophisticated methods in literature [2], [4], [11], [12], [14], [17] ranging from high-gain observer schemes over integration-based methods to the use of splines for the reconstruction. A closer look revealed that to date there is little investigation performed that compares all or a subset of methods in regard to their use in industrial applications. To this end, this article takes a closer look at three different methods a weighted-time integration-based differentiator [14], [17], a B-spline based differentiator [10] and a method based on sliding-mode observers [12]. All methods are compared w.r.t. their noise suppression capabilities, their computational complexity and their estimation accuracy, of course. This will help to uncover their main (dis)advantages and leads to an application dependent recommendation for two of the three methods.

The set-up for comparison is depicted in Fig. 1. This directly leads to a description of the problem as follows: Given a noise corrupted signal $x(t)$ estimate a "noise-free" signal $\hat{x}(t)$ and its derivatives $\dot{\hat{x}}(t), \ddot{\hat{x}}(t), \ldots, \hat{x}^{(n)}(t)$. All methods are compared using a generally unknown but deterministic reference or base signal that is corrupted with measurement noise. Each differentiator then relies on the use of a finite amount of discrete measurements over its own observation windows. From this, derivatives up to the fifth order are computed and compared. In order to further highlight industrial applicability all methods are also compared on a typical trajectory observed in motion control systems. Firstly, such systems often face an additional quantization noise. Secondly, many jerk-limited motion trajectories are non-smooth in one or the other variable. This marks a severe difficulty for the estimation process.

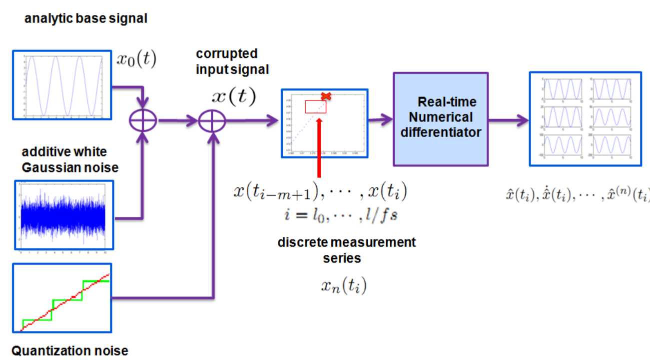


Fig. 1. Task overview of a real-time numerical differentiator

The remainder of this article is organized as follows: The next section starts with the mathematical exposition of the respective integration algorithms under study. Then, Section III contains the comparison and discussion of the methods using two specific test cases – a motion control trajectory and the reconstruction of a chaotic signal. Finally, a recommendation and conclusion is provided in Section IV.

Generally, all material presented here must be seen as a snapshot of a much broader set of results contained in [20].

II. Methods for Differentiation

As aforementioned, in this section three different methods will be presented for estimating the derivative of a noisy measurement signal. All three methods apply different principles for obtaining the derivative which is one reason for the choice made.

A. Weighted-time Integration Differentiator

The first method to be introduced is the weighted-time integration-based differentiator presented in [14], [17]. Simply put, the method estimates the derivative of a signal by

Kim D. Listmann is with the ABB AG, Corporate Research Center Germany, Wallstadter Str. 59, 68526 Ladenburg, GERMANY. kdl@ieee.org, Zheng Zhao is with the Conti Temic Microelectronic GmbH, Sieboldstr. 19, 90411 Nürnberg. zheng.zhao@continental-corporation.com.

978-3-952-41734-8/©2013 EUCA

approximating the signal using a particular Taylor polynomial of sufficient order, identifying the coefficients of this polynomial as the derivatives to be estimated and performing a numerical integration over a predefined time interval to obtain these coefficients. Then noise suppression comes as a byproduct due to the numerical integration (smoothing) used.

As a starting point, every analytic signal can be rewritten to a full Taylor series. For approximating its behavior this infinte series can be cut after a finite amount of elements, where the number of elements directly correlates to the accuracy of the approximation. Evaluating the Taylor series around the development point zero (typically known as Maclaurin series) and taking $\tau = -t \in [0, T]$ as the independent variable, the approximated signal reads

$$x(-\tau) \approx \sum_{i=0}^{n} \frac{a_i}{i!}(-\tau)^i.$$

Thus, through $x(t) = x(-\tau)$ we have $x(t) \approx \sum_{i=0}^{n} \frac{b_i}{i!} t^i$. Then, the n-th derivative at $t = t_i$ can be calculated as

$$\hat{x}^{(n)}(t_i) = \hat{x}^{(n)}(0) = b_n = (-1)^n a_n(-\tau), \tag{1}$$

where a_n is computed based on the noisy signal $x(t_i - \tau)$ over $\tau \in [0, T]$. To this end, compute the Laplace transform of (1)

$$\tilde{x}(s) = \sum_{j=0}^{n} \frac{a_{j,n}}{s^{j+1}}.$$

Then, multiplication with s^{n+1} yields

$$s^{n+1}\tilde{x}(s) = \sum_{j=0}^{n} a_{j,n} s^{n-j}, \tag{2}$$

where $n = 0, 1, 2, \cdots, N$. Next, the Maclaurin coefficients can be computed successively starting with $a_{0,n}$. Mathematically this leads to a triangle system of equations formed by taking the i-th derivative of (2) with $i \in \{0, 1, \ldots, N\}$, respectively. This can be computed to

$$\frac{d^i}{ds^i}(s^{n+1}\tilde{x}(s)) = \sum_{j=0}^{n-i} a_{j,n} \frac{(n-j)!}{(n-j-i)!} s^{n-j-i}.$$

and the application of Leibniz' rule for partial derivatives leads to the simplified equation

$$\sum_{j=0}^{n-i} a_{j,n} \frac{(n-j)!}{(n-j-i)!} s^{n-j-i} = \sum_{k=0}^{i} \binom{i}{k} \frac{(n+1)!}{(n+1-k)!} s^{n+1-k} \tilde{x}^{(i-k)}(s), \tag{3}$$

where $\binom{i}{k} = \frac{i!}{k!(i-k)!}$, $i = 0, \cdots, n$ and $n = 0, 1, 2, \cdots, N$. In order to purely perform integration and introduce the least possible delay, (3) must be multiplied with s^{n-2} to obtain

$$\sum_{k=0}^{i} \binom{i}{k} \frac{(n+1)!}{(n+1-k)!} s^{-1-k} \tilde{x}^{(i-k)}(s) = \sum_{j=0}^{n-i} a_{j,n} \frac{(n-j)!}{(n-j-i)!} s^{-j-i-2}.$$

Retransformation into the time domain over the predefined interval $[0, T]$ gives

$$\sum_{k=0}^{i} \binom{i}{k} \frac{(n+1)!}{(n+1-k)!} \int_0^{T_n} \int_0^{\alpha_1} \cdots \int_0^{\alpha_k} (-\alpha_k)^{i-k} \tilde{x}(\alpha_{k+1}) d\alpha_{k+1} d\alpha_k \cdots d\alpha_1$$
$$= \sum_{j=0}^{n-i} a_{j,n} \frac{(n-j)!}{(n-j-i)!} \frac{T^{i+j+1}}{(i+j+1)!}. \tag{4}$$

Consequently, by applying the Cauchy formula [8]

$$\int_0^{T} \int_0^{\alpha_1} \cdots \int_0^{\alpha_{k-1}} x(\alpha_k) d\alpha_k d\alpha_{k-1} \cdots d\alpha_1 = \int_0^{T} \frac{1}{(k-1)!} (T-t)^{(k-1)} x(t) dt,$$

(4) can be simplified to

$$\sum_{k=0}^{i} \binom{i}{k} \frac{(n+1)!}{(n+1-k)!} \int_0^{T_n} \frac{1}{k!} (T_n - t)^k (-t)^{i-k} \tilde{x}(t) dt$$
$$= \sum_{j=0}^{n-i} a_{j,n} \frac{(n-j)!}{(n-j-i)!} \frac{T_n^{i+j+1}}{(i+j+1)!},$$

with $i = 0, \cdots, n$ and $n = 0, 1, 2, \cdots, N$. Finally, to compute the coefficients we need to solve

$$P \begin{bmatrix} a_{0,n} \\ a_{1,n} \\ \vdots \\ a_{n,n} \end{bmatrix} = \begin{bmatrix} \int_0^{T_n} q_1(t)\tilde{x}(t)dt \\ \int_0^{T_n} q_2(t)\tilde{x}(t)dt \\ \vdots \\ \int_0^{T_n} q_n(t)\tilde{x}(t)dt \end{bmatrix} \approx \begin{bmatrix} \int_0^{T_n} q_1(t)x(t_i - t)dt \\ \int_0^{T_n} q_2(t)x(t_i - t)dt \\ \vdots \\ \int_0^{T_n} q_n(t)x(t_i - t)dt \end{bmatrix},$$

with

$$q_i(t) = \sum_{k=0}^{i-1} \frac{(i-1)!(n+1)!}{(k!)^2(i-k-1)!(n+1-k)!} (T_n - t)^k (-t)^{i-k-1},$$
$$P_{i,j} = (-1)^{j-1} \frac{(n-j+1)!}{(n-j-i+2)!} \frac{T_n^{j+i-1}}{(j+i-1)!}.$$

Here, T_n represents the length of the observation window, $i = 1, \ldots, n+1$, $j = 1, \ldots, n-i+2$ and $n = 0, 1, 2, \ldots, N$. The analytic expression of $a_{n,n}$ with $n = 0, 1, 2, \ldots, N$ can be summarized through a system of equations, where each $a_{n,n}$ for the current time point $t = t_i$ is a weighted integration of the noisy signal within the sliding observation window. Note that this can be further simplified using an approach from [15], so that the differentiator reduces to an FIR filter. Numerical cost become minor then.

As aforementioned analytic expressions of the estimates $\hat{x}(t_i), \dot{\hat{x}}(t_i), \cdots, \hat{x}^{(N)}(t_i)$ for $t = t_i$ can be obtained based on the time domain expressions of $a_{n,n}$ with $n = 0, 1, 2, \cdots, N$, respectively by

$$\hat{x}^{(n)}(t_i) = (-1)^n a_{n,n}.$$

At this point it is worth mentioning that the presented differentiator inherently fulfills a least-squares error condition, which was already shown in [17]. Further, we would like to point out, that this is the first time that a Maclaurin series was used for the representation, whereas in [17], [14] a general Taylor approximation is used. The Maclaurin series used here, further simplifies the general approach.

B. B-Spline Differentiator

The next and second method to be presented relies on fitting a curve, namely a B-spline, into the given set of measurements smoothly [10]. Since the curve is a spline, differentiation can then be performed analytically. Fitting the spline to the measurements, an optimization is performed taking the partitioning of the spline curves and the error of the spline curve w.r.t. the measurements into account.

At the current time point $t = t_i$, the approximated B-spline of order $2n$ (suggested in [10]) over the moving observation window $t_{i-m+1} \leq t \leq t_i$ can be expressed as a linear combination of B-spline control points α_j as

$$\tilde{x}(t) = \sum_{j=0}^{m-1} \alpha_j B_{j,2n}(t_{i+j-m+1}).$$

Here, $B_{j,2n}$ is a basis function of order $2n$. It is in the form of a weighted sum of control points. If the B-spline is of order $2n$, then its value within $t_{i-m+1} \leq t \leq t_i$ is a weighted sum of at most $2n$ control points out of m control points, i.e., at most $2n$ control points are active at the same time. In addition, the sum of basis functions at a certain time $t \in [u_j, u_{j+1}]$ is one. Hence, the recursive computation scheme of basis functions can be expressed mathematically as

$$B_{j,1} = \begin{cases} 1 & \text{if } u_j \leq t_{i+j-m+1} < u_{j+1} \\ 0 & \text{otherwise,} \end{cases}$$

with $j = 0, 1, \cdots, m + 2n - 2$ and

$$\begin{aligned} B_{j,k}(t) = & \frac{t_{i+j-m+1} - u_j}{u_{j+k-1} - u_j} B_{j,k-1}(t_{i+j-m+1}) \\ & + \frac{u_{j+k} - t_{i+j-m+1}}{u_{j+k} - u_{j+1}} B_{j+1,k-1}(t_{i+j-m+1}), \end{aligned}$$

with $k = 2, 3, \cdots, 2n$, $j = 0, 1, \cdots, m - 1$. If two knots u_j are identical, any resulting indeterminate forms, e.g., $\frac{0}{0}$, are deemed to be 0.

The exact value of a basis function of an order larger than one is determined by recurrence relations depending on the ratio of the distance between the evaluated time instant $t_{i+j-m+1}$ to the knot u_j and $\|u_{j+k} - u_{j+1}\|$ along with the one of that between the evaluated time point to the knot u_{j+k} and $\|u_{j+k-1} - u_j\|$. According to [3], given m measurements $t_{i-m+1}, t_{i-m+2}, \cdots, t_i$ for a B-spline of order $2n$ these knots need to be computed as

$$\begin{aligned} u_0 &= u_1 = \cdots = u_{2n-1} = t_{i-m+1} \\ u_{l+2n-1} &= \frac{1}{2n-1} \sum_{j=l+i-m+1}^{l+2n-m-1} t_j, \text{ for } l = 1, 2, \ldots, m - 2n \\ u_m &= u_{m+1} = \cdots = u_{m+2n-1} = t_i. \end{aligned}$$

Moreover, it was shown in [3] that the derivatives up to the $(2n - 2)$-th order of a $2n$ order B-spline can easily be obtained by a difference scheme. Thus, obtaining derivatives up to the order n from the reconstructed B-spline is possible. The p-th derivative is given by

$$D^p(\sum_j a_j B_{j,2n}) = \sum_j a_j^{(p+1)} B_{j,2n-p},$$

with

$$a_j^{(p+1)} := \begin{cases} a_j & \text{if} p = 0, \\ \frac{a_j^{(p)} - a_{j-1}^{(p)}}{(t_{i+j-m+1+2n-p} - t_{i+j-m+1})/(2n-p)} & \text{otherwise.} \end{cases}$$

Here $p = 0, 1, 2 \cdots, 2n$. The computed p-th derivative exists only within $t_{i-m+p+1} \leq t \leq t_i$, due to the value of that at the first p measurement points being not available anymore. In addition, if the derivatives of a B-spline at a certain time point or within a fixed interval will be determined, the calculation domain can be further refined.

At this point the only thing missing is the value of the control points α_j. These will be obtained as the optimal solution of the minimization problem

$$\min \xi = \frac{1}{m}(X - B\alpha)^T (X - B\alpha) + \lambda \alpha^T B^T R B \alpha. \tag{5}$$

Here, $R = H^T H$, with the $(m - n) \times m)$ matrix H defined by

$$H_{i,j} = \begin{cases} (-1)^{n+j-i} \frac{n!}{(j-i)!(n-j+i)!} & i \leq j \leq i + n \leq m, \\ 0 & 1 \leq j < i \text{ or } i + n < j \leq m, \end{cases}$$

with $i \in \{1, \ldots, m - n\}$. B is the B-spline basis function matrix, α is the control point vector and X is the matrix of measurements. The so called regularization parameter λ can also be considered a free variable.

For ξ to be a minimum, the conditions $d\xi/d\alpha \stackrel{!}{=} 0$ and $d^2\xi/d\alpha^2 \stackrel{!}{\geq} 0$ need to be fulfilled [16]. Given (5), using matrix calculus [5], the optimal control vector $\alpha^\star$ is given by

$$\alpha^\star = (m\lambda RB + B)^{-1} X.$$

if the condition

$$\frac{2}{m} B^T B + 2\lambda B^T R B \stackrel{!}{\geq} 0$$

holds. In this set-up λ is computed numerically using a Newton-Raphson method. Following [10], this results from solving the generalized cross validation criterion [18], [19]

$$\min L(\lambda) = \frac{\frac{1}{m}\|m\lambda R(I + m\lambda R)^{-1} X\|^2}{[\frac{1}{m}\text{trace}(m\lambda R(I + m\lambda R)^{-1})]^2}$$

whose computation is omitted due to space restrictions.

C. Sliding-mode Differentiator

The last method to be presented in this comparison is drawn from the category of sliding mode observers. The basic idea is the following: Successively apply a method developed in [11] for estimating the first derivative of a measurement signal. To this end, let $x(t)$ be a locally bounded, measurable, noisy input signal with base signal $x_0(t)$. Then, the estimated base signal $z_0 := \hat{x}_0(t)$ and its first derivative $z_1 := \hat{\dot{x}}_0(t)$ can be obtained by

$$\begin{aligned} \dot{z}_0 &= -\lambda_0 |z_0 - x|^{1/2} \text{sign}(z_0 - x) + z_1, \\ \dot{z}_1 &= -\lambda_1 \text{sign}(z_1 - \dot{z}_0), \end{aligned} \tag{6}$$

with initial conditions $z_0(0) = x(0)$, $z_1(0) = 0$, and $\lambda_0, \lambda_1 > 0$. The introduction of the variable z_0 and z_1 is to distinguish the difference between the derivative of the estimated base signal $\frac{d}{dt}\hat{x}_0(t) \equiv \dot{z}_0$ and the estimated first derivative $\hat{\dot{x}}_0(t) \equiv z_1$. This scheme results in discontinuous behavior of the estimated quantities, but theoretically reconstructs the true unknown signal (if it is noise free and sampled infinitely-frequent) in finite time. Even if a noisy (bounded) signal is estimated, it provides the best possible error asymptotics [12].

The extension to a differentiator of arbitrary order makes use of the fact that the estimated first derivative has a bounded error. Thus, it can also be regarded as a noisy input signal with bounded noise for the estimation of the second derivative and so on. Consequently, an arbitrary order differentiator with a recursive structure of the 2-sliding differentiator in (6) is reasonable. The arbitrary order sliding-mode based differentiator proposed in [12] is again slightly modified by using an adjusted Lipschitz constant. It can estimate a base signal and the derivatives up to the n-th order. These are denoted $z_0 := \hat{x}(t), z_1 := \hat{\dot{x}}(t), \cdots, z_n := \hat{x}^{(n)}(t)$ and estimated through

$$\begin{aligned}
\dot{z}_0 &= -\lambda_0 \hat{L}^{1/(n+1)} |z_0 - x|^{n/(n+1)} \operatorname{sign}(z_0 - x) + z_1, \\
\dot{z}_1 &= -\lambda_1 \hat{L}^{1/n} |z_1 - \dot{z}_0|^{(n-1)/n} \operatorname{sign}(z_1 - \dot{z}_0) + z_2, \\
&\vdots \\
\dot{z}_i &= -\lambda_i \hat{L}^{1/(n-i+1)} |z_i - \dot{z}_{i-1}|^{(n-i)/(n-i+1)} \operatorname{sign}(z_i - \dot{z}_{i-1}) + z_{i+1}, \\
&\vdots \\
\dot{z}_{n-1} &= -\lambda_{n-1} \sqrt{\hat{L}} |z_{n-1} - \dot{z}_{n-2}|^{1/2} \operatorname{sign}(z_{n-1} - \dot{z}_{n-2}) + z_n, \\
\dot{z}_n &= -\lambda_n \hat{L} \operatorname{sign}(z_n - \dot{z}_{n-1}),
\end{aligned} \tag{7}$$

where $\lambda_i > 0$ for $i \in \{0, \ldots, n\}$ must be chosen properly.

Since the sliding-mode based differentiator is a successive scheme, once the parameters $\lambda_0, \lambda_1, \cdots, \lambda_{n-1}$ along with an adaption scheme for the Lipschitz constant $\hat{L}$ are chosen properly for an $(n-1)$-th order differentiator, only one parameter, λ_n, needs to be tuned with a corresponding adaption scheme of $\hat{L}$. A recommended parameterization [12] for a fifth-order differentiator is $\lambda_5 = 1.1$, $\lambda_4 = 1.5$, $\lambda_3 = 3$, $\lambda_2 = 5$, $\lambda_1 = 8$ and $\lambda_0 = 12$. Moreover, the convergence process of each estimated derivative is faster the larger the chosen Lipschitz constant is as long as the sampling frequency is sufficiently high.

The accuracy and convergence speed of (7) highly relies on a good adjustment of the estimated Lipschitz constant $\hat{L}$ w.r.t. the true Lipschitz constant L of the signal. Recall that here the estimated Lipschitz $\hat{L}$ constant denotes the highest slope of the $(n+1)$-th derivative within a certain time interval. Practically, the assigned Lipschitz constant should be adjusted based on the Lipschitz constant of the n-th derivative of the base signal, in order to regulate the amplitude of the switched gains. Switched gains are essential factors affecting the error bounds and the convergence speed of this differentiator.

After introducing $\delta_i := z_i - x_0^{(i)}(t) = z_0^i - x_0^{(i)}(t)$ with $i = 0, 1, \cdots, n$, the mathematical expression of n-th order differentiator (7) can be rewritten to

$$\begin{aligned}
\dot{\delta}_0 =& -\lambda_0 \hat{L}^{1/n+1} |\delta_0 - x_n(t)|^{n/n+1} \operatorname{sign}|\delta_0 - x_n(t)| + \delta_1, \\
\dot{\delta}_1 =& -\lambda_1 \lambda_0^{n-1/n} \hat{L}^{2/n+1} |\delta_0 - x_n(t)|^{n-1/n+1} \operatorname{sign}\, n|\delta_0 - x_n(t)| + \delta_2, \\
&\vdots \\
\dot{\delta}_{n-1} =& -\lambda_{n-1} \lambda_{n-2}^{1/2} \cdots \lambda_1^{1/n-1} \lambda_0^{1/n} \hat{L}^{n/n+1} |\delta_0 \\
& - x_n(t)|^{1/n+1} \operatorname{sign}|\delta_0 - x_n(t)| + \delta_n, \\
\dot{\delta}_n =& -\lambda_n \hat{L} \operatorname{sign}|\delta_0 - x_n(t)|.
\end{aligned} \tag{8}$$

As seen in (8), the Lipschitz constant $\hat{L}$ constructs the amplitudes of switched gains along with the parameters λ and the distance between the estimated base signal and the input signal $|z_0 - x(t)| = |\delta_0 - x_n(t)|$. Thus, the applied Lipschitz constant can adjust the convergence speed of this differentiator [11]. A larger Lipschitz constant results in quicker convergence. However, it may also result in increased chattering. Moreover, because the best possible error of the estimated i-th derivative of an n-th order differentiator is proportional to $\hat{L}^{i/(n+1)} \epsilon^{(n+1-i)/(n+1)}$ with $i = 0, 1, \cdots, n$ [12], the accuracy of the best possible estimation after certain transient time is increased with a smaller Lipschitz constant. So the Lipschitz constant trades-off estimation error and speed of convergence in this setting. In principle, there is two possibilities: (i) If a globally valid Lipschitz constant estimate is used in an n-th order differentiator, the asymptotic convergence speed will theoretically remain the same over the whole observation period – for the quickly as well as for the slowly varying parts of the signal. One the one hand, this prevents the algorithm from achieving high accuracy of the estimates for the parts with smaller Lipschitz constant (whose best accuracy may be attainable with much smaller switching gains). On the other hand, although a local Lipschitz constant may be used to generate better estimation locally, it can cause different asymptotic convergence speed in the whole implementation and thus, results in time-varying delay in the estimates. (ii) Adapt $\hat{L}$ to respect local changes in the signal. This might however lead to unsatisfactory if not unstable estimation dynamics.

III. Comparison of the Methods

After introducing the different methods it's time to compare their properties. To this end, we decided to focus on their computational complexity, maintainability/robustness, and noise suppression abilities. This will be done based on two different dynamics to be estimated: (i) a Lorenz attractor as an example for a quickly varying, aperiodic signal and (ii) a jerk limited motion trajectory as a standard signal in practical applications.

A. Computational Complexity

Looking back at the methods, the complexity for the implementation differs greatly. There is the B-spline based integrator that relies on the solution of two optimization problems vs. the other two methods that mainly need numerical integration schemes to obtain their estimates. Moreover, the order for a B-spline polynomial was at best already twice as high as the order of the Maclaurin series for the weighted-time integration-based differentiator.

To differentiate between the sliding-mode differentiator and the integration-based differentiator keep in mind that for the latter, a system of linear equations must finally be solved to obtain the coefficients. For the former the implementation is very slim since only discontinuous dynamics are handled. Note that in both cases the integration using a forward Euler method [1] was sufficient.

As a consequence, the "cheapest" integrator from a computational point of view is the sliding-mode differentiator, followed by the weighted-time integration-based method. If we use the FIR implementation of [15] for the latter this order will swap, of course.

B. *Maintainability*

This picture changes when looking at the usability for unexperienced people out there. The main motivation for evaluating this is because at the end, one differentiator will estimate sensor signals in a technical device. This implies that people have to check the output of the algorithm and probably also tune it to a certain extent. Then, the sliding-mode differentiator certainly is the weakest of all. It's dependency on the Lipschitz constant of the signal is shear inscrutable and there is quite some tuning effort required to get from a first implementation to the final status. Here, the advantage of an optimization scheme is apparent: The algorithm does everything it needs and the user does not have to set-up or observe a single variable. But also the integration-based method is good to maintain and use. It contains literally no tuning possibilities and simply performs an integration and the solution of a linear system of equations.

C. *Chaotic Dynamics*

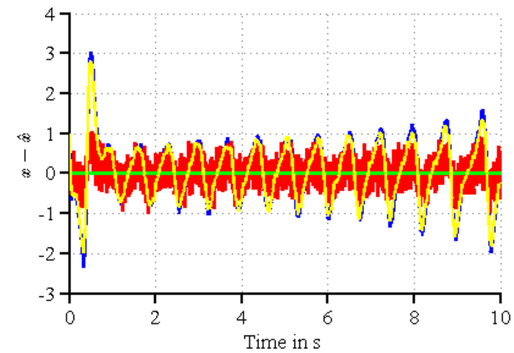
Results for the estimation of a Lorenz attractor [13] are provided in Fig. 2 including all three differentiators and a finite difference scheme combined with a low pass filter (2nd order Butterworth, cut-off freq. 100 Hz) as a baseline for comparison. Note that the estimation error is depicted here. For all methods the reconstructed base signal (SNR 25 dB, sampling@1000 Hz), first and second derivative are displayed. The Lipschitz constant for the sliding-mode differentiator was set to $3 \cdot 10^9$.

Obviously, the sliding mode-based differentiator performs very well for this signal. It possesses superior noise suppression quality and the time delay introduced into the signal is minimal. The B-spline differentiator is in this case not capable of estimating the first derivative appropriately already. The standard appraoch using a finite difference and low-pass filter is competitive only if just the first derivative is compared. All higher derivatives suffer from noise contained in the signal.

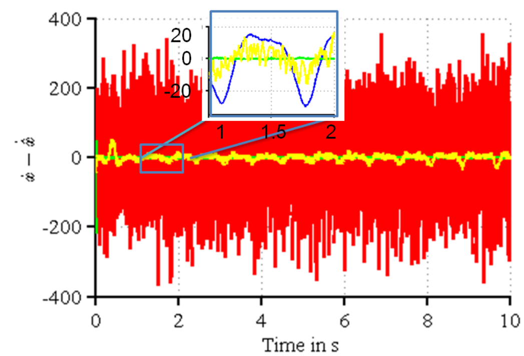
D. *Motion Control Trajectories*

Finally, we like to compare the capabilities of all the methods in the estimation of typical trajectories for motion systems. Based on a jerk-limited trajectory (maximum jerk 0.5 m/s^3 the differentiators should estimate the acceleration, the velocity and the position accurately (SNR 45 dB, sampling@1000 Hz). Fig. 3 shows the results obtained. In this setting 3.5 was used as Lipschitz constant for the sliding-mode differentiator.

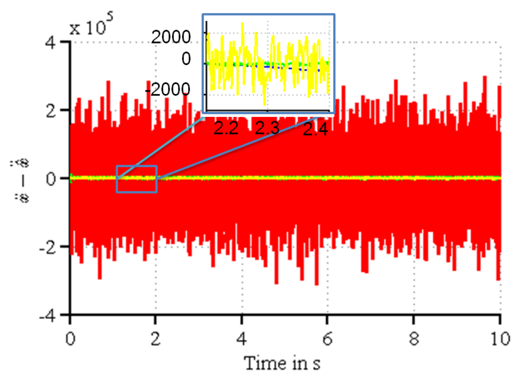
This time the sliding-mode based differentiator and the weighted-time integration based method are head to head. The former has a slight advantage reconstructing the position, while the latter performs better on estimating the derivatives. As before the B-spline based differentiator is not really able



(a) Estimation error base signal

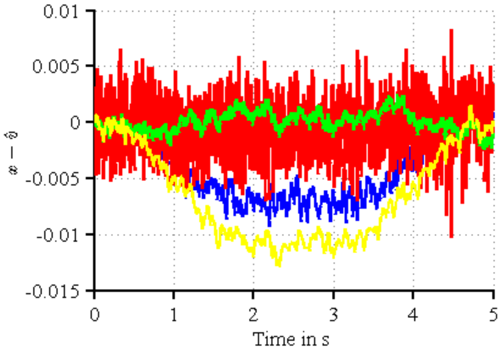


(b) Estimation error first derivative

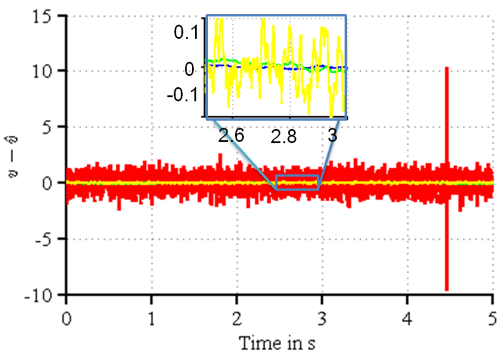


(c) Estimation error second derivative

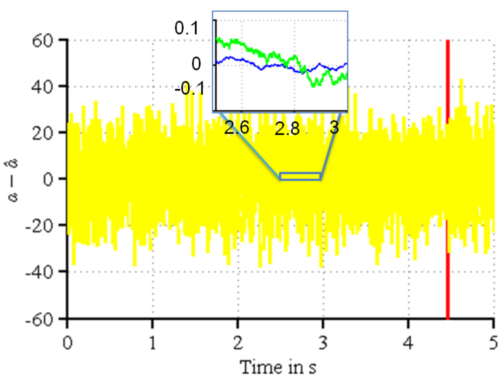
Fig. 2. Estimation results for the chaotic process comparing the estimation error for all three differentiators (weighted time = blue, B-spline = red, sliding mode = green) and the finite-difference low pass combination (yellow).



(a) Estimation error position trajectory



(b) Estimation error velocity



(c) Estimation error acceleration

Fig. 3. Estimation results for a standard motion profile (jerk limited) comparing the estimation error for all three differentiators (weighted time = blue, B-spline = red, sliding mode = green) and the finite-difference low pass combination (yellow).

to produce reliable estimates for the given measurements. The standard approach again stays competitive up to the first derivative but may not be used for derivatives of higher order.

IV. Conclusion

This article reviewed three different methods for the estimation of derivatives from noisy measurements. the first was based on approximating the signal as a Maclaurin series and obtaining the derivatives by numerical integration. The second was based on fitting the curve of a B-spline to a set of data points in an optimal way. Last but not least, a sliding-mode differentiator was used that lead to a discontinuous system when closing the loop.

The main attributes the methods were compared on were their simplicity, their noise suppression quality and accuracy of estimation. This was done for a chaotic attractor as well as for a classical motion control trajectory. Refering to a stanard set-up, the comparison included a finite difference scheme with 2nd-order Butterworth low-pass filter.

Providing a recommendation depends on the considered signal: If one knows the trajectories of the system very well and it is not too fast, then the sliding-mode differentiator proves beneficial if a high sampling frequency is applicable. This is mainly because it has nice convergence properties and results from a very slim implementation. If on the other hand, the above does not hold and the signals are rather unknown, the weighted time integration-based differentiator is the primary choice. It has superior noise rejection properties and gives the best available compromise. Moreover, it outperforms all others for the higher order derivatives (second derivative and above).

References

[1] J. C. Butcher. *Numerical Methods for Ordinary Differential Equations*. John Wiley & Sons, 2003.

[2] Y. Chitour. Time-varying high-gain observers for numerical differentiation. *IEEE Trans. Automat. Control*, 47:1565–1569, 2002.

[3] C. de Boor. *A Practical Guide to Splines*. Springer, 2001.

[4] F. Deza. High gain estimation for nonlinear systems. *Systems & Control Letters*, 18(4):295 – 299, 1992.

[5] C. Felippa. *Introduction to Finite Element Methods*. University of Colorado, 2003.

[6] M. Fliess and C. Join. Model-free control and intelligent pid controllers: towards a possible trivialization of nonlinear control? [Online], 2009. arXiv:0904.0322.

[7] M. Fliess, C. Join, and H. Sira-Ramirez. Non-linear estimation is easy. [Online], 2007. arXiv:0710.4486.

[8] G. B. Folland. *Advanced Calculus*. Prentice Hall, 2002.

[9] Z.-S. Hou and J.-X. Xu. On data-driven control theory: the state of the art and perspective. *Acta Automatica Sinica*, 35(6):650–667, 2009.

[10] S. Ibrir. A numerical procedure for filtering and efficient high-order signal differentiation. *Int. J. Appl. Math. Comput. Sci*, 14(2):201–208, 2004.

[11] A. Levant. Robust exact differentiation via sliding mode technique. *Automatica*, 34(3):379–384, 1998a.

[12] A. Levant. Higher-order sliding modes, differentiation and output-feedback control. *Int.J.Control*, 76(9/10):924–941, 2003.

[13] E. N. Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2):130–141, 1963.

[14] 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*, pages 1–19, 2007.

[15] M. Mboup, C. Join, and M. Fliess. Numerical differentiation with annihilators in noisy environment. *Numerical Algorithms*, 50(4):439–467, 2009.

[16] J. Nocedal and S. Wright. *Numerical Optimization*. Springer, 2nd edition, 2000.

[17] C. Noethen. Beitraege zur Rekonstruktion nicht direkt gemessener Groessen bei der Silizium-Einkristallzuechtung nach dem Czochralski-Verfahren. Master's thesis, TU Dresden, Dresden, Germany, 2007.

[18] C. Reinsch. Smoothing by spline functions. *Numer. Math.*, 10:177–183, 1967.

[19] G. Wahba. Smoothing noisy data with spline functions: Estimation the correct degree of smoothing by the method of generalized cross-validation. *Numer. Math.*, 18(3):377–403, 1979.

[20] Z. Zhao. Reconstruction of derivatives from noisy measurements - an optimization and comparison. Master's thesis, RWTH Aachen, Aachen, Germany, 2012.