Optimal robust exact differentiation via linear adaptive techniques☆

Richard Seeber [a],[*], Hernan Haimovich [b]

[a] *Graz University of Technology, Institute of Automation and Control, Christian Doppler Laboratory for Model Based Control of Complex Test Bed Systems, Graz, Austria*
[b] *Centro Internacional Franco-Argentino de Ciencias de la Información y de Sistemas (CIFASIS) CONICET-UNR, 2000 Rosario, Argentina*

ARTICLE INFO

ABSTRACT

The problem of differentiating a function with bounded second derivative in the presence of bounded measurement noise is considered in both continuous-time and sampled-data settings. Fundamental performance limitations of causal differentiators, in terms of the smallest achievable worst-case differentiation error, are shown. A robust exact differentiator is then constructed via the adaptation of a single parameter of a linear differentiator. It is demonstrated that the resulting differentiator exhibits a combination of properties that outperforms existing continuous-time differentiators: it is robust with respect to noise, it *instantaneously* converges to the exact derivative in the absence of noise, and it attains the smallest possible—hence optimal—upper bound on its differentiation error under noisy measurements. For sample-based differentiators, the concept of quasi-exactness is introduced to classify differentiators that achieve the lowest possible worst-case error based on sampled measurements in the absence of noise. A straightforward sample-based implementation of the proposed linear adaptive continuous-time differentiator is shown to achieve quasi-exactness after a single sampling step as well as a theoretically optimal differentiation error bound that, in addition, converges to the continuous-time optimal one as the sampling period becomes arbitrarily small. A numerical simulation illustrates the presented formal results.

1. Introduction

The signal differentiation problem consists in estimating a signal's derivatives based on signal measurements. This seemingly simple task becomes complicated in the presence of measurement noise and other perturbations, even if measurements are assumed to be available continuously over time. In a practical setting, moreover, measurements become available only at sampling instants, and calculations need to be performed on a digital computer; this may present additional obstacles to differentiation.

Methods for signal differentiation include algebraic methods involving elementary differential algebraic (linear) operations (Mboup et al., 2009), linear high-gain observers (Khalil &

Praly, 2014; Vasiljevic & Khalil, 2008) and sliding-mode differentiators (Levant, 1998, 2003; Yu & Xu, 1996). The linear differentiators can have good measurement noise rejection capability but are not exact, meaning that in the absence of noise their output is not ensured to converge to the true value of the signal derivative. Sliding-mode differentiators, in contrast, can be exact but this may lead to the exact differentiation also of some (differentiable) noise signals. Some strategies even aim to combine the positive features of linear and sliding-mode differentiators (Ghanes et al., 2020).

One specific differentiation problem of particular interest in relation to sliding-mode control is the estimation of a signal's first n derivatives having knowledge of a bound for the $n + 1$th derivative. In this context, the presence of measurement noise limits the accuracy of any differentiator with a bound depending on noise amplitude and known derivative bound. The expressions for these accuracy limitations are related to the so-called Landau–Kolmogorov inequalities (Kolmogorov, 1962; Schoenberg & Cavaretta, 1970), cf. also Levant et al. (2017). The differentiation accuracy for many sliding-mode differentiators can be shown to be of the asymptotic order indicated by these inequalities as the noise amplitude tends to zero (Cruz-Zavala et al., 2011; Levant, 2003).

Since the derivative estimates are usually required for achieving some subsequent control objective, an important feature of

* Corresponding author.
E-mail addresses: richard.seeber@tugraz.at (R. Seeber), haimovich@cifasis-conicet.gov.ar (H. Haimovich).

any differentiator is its convergence speed. Sliding-mode differentiators can converge in finite time (Levant, 1998), whereas linear differentiators only do so asymptotically (Vasiljevic & Khalil, 2008). Moreover, some sliding-mode differentiators can even converge in a finite time that is independent of initial conditions (Cruz-Zavala et al., 2011), a situation that is called fixed-time convergence (Bhat & Bernstein, 2000). In addition, such differentiators can be designed to achieve any given bound on the convergence time (Seeber et al., 2021).

When only sampled measurements are available and differentiator implementation becomes digital, then the differentiators conceived as continuous-time systems must be discretized. Care must be taken, however, because with improper discretization the accuracy of the resulting discretized differentiator can worsen significantly compared to the continuous-time one (Levant, 2012). The analysis of different discretization techniques becomes then highly relevant (Andritsch et al., 2021; Carvajal-Rubio et al., 2021; Mojallizadeh et al., 2021b). Livne and Levant (2014) propose a proper discretization that preserves the accuracy asymptotics of a continuous-time differentiator. Several extensions of this idea are given by Barbot et al. (2020). Moreover, Hanan et al. (2021) introduce a rather slight modification that allows to lower the resulting discrete differentiator's output chattering in the absence of noise. Filtering differentiators (Levant & Livne, 2020; Levant & Yu, 2018) are even capable of filtering out unbounded noise with small average values while preserving all advantageous features of the standard sliding-mode-based differentiators. A radically different strategy for signal differentiation in a digital implementation setting is to directly consider the information carried by noisy sampled measurements accounting for noise magnitude and known derivative bound. If a suitable bound on the noise magnitude is known, then the differentiation problem can be performed through solving specific convex optimization problems (Haimovich et al., 2022) in the form of linear programs. This strategy is shown to achieve the best possible worst-case accuracy and to have explicitly computable fixed-time convergence and accuracy bounds for first-order differentiation, provided that the noise amplitude is known.

For the continuous-time case, an interesting consequence of the results in Haimovich et al. (2022) is that the best possible worst-case accuracy may also be achieved by a linear finite difference. However, this is true only if the noise magnitude is known to the differentiator; if the actual noise magnitude happens to be lower than the assumed bound, then performance can worsen greatly. In particular, such a linear differentiator can never be exact for arbitrary signals with bounded second derivative. The use of such finite differences for robust output-feedback control was explored by Levant (2007), who showed that when utilized in conjunction with homogeneous sliding-mode controllers, differentiation via finite differences turns into an appealing strategy by adaptation of the sampling time. This adaptation allowed to achieve similar asymptotic accuracy as is obtained when using sliding-mode differentiators. However, its use is restricted to differentiation in a closed feedback loop.

In this paper, first-order differentiation of noisy signals is considered. Via adaptation of the parameter of a linear finite difference, a robust exact differentiator is constructed that achieves *optimal differentiation accuracy*. Contrary to what is called asymptotic optimality in Levant (1998), which relates just to optimality of the exponents in the accuracy expression, optimality is achieved here in the sense that the differentiation error is ultimately bounded from above by the *smallest possible upper bound*. It is further shown that the convergence time of the resulting differentiator in the absence of noise is zero, i.e., that it outputs the true derivative from the beginning, and hence converges faster than any fixed-time differentiator. For practical realization, it is

shown that a straightforward sample-based realization achieves the lowest possible worst-case error among all sample-based differentiators in the absence of noise—a fact that is formalized by introducing the notion of quasi-exactness for sample-based differentiators. Moreover, the proposed realization is shown to retain its optimal convergence behavior and its optimal accuracy with respect to the noise.

After a brief problem statement, Section 2 introduces formal definitions for worst-case differentiation error, exactness, and robustness that take the initial values of the signal into account. Sections 3, 4, and 5 then consider the continuous-time case, while Sections 6 and 7 deal with the case of sampled measurements. Sections 3 and 6, in particular, show lower bounds for the worst-case error of all (causal) differentiators; the former section discusses such performance limits for arbitrary and for exact differentiators, which motivate the notion of (optimal) differentiation accuracy, while the latter section shows performance limitations incurred by sampling, leading to the notion of quasi-exactness of sample-based differentiators. As a preliminary result for the considered linear adaptive strategy, Section 4 shows that a finite difference is an optimal causal linear differentiator, which is not exact, however. The main results—the proposed optimal robust exact differentiators—are then presented in Sections 5 and 7 for the continuous-time and the sampled-data case, respectively; their formal properties, in particular, are summarized in the main Theorems 5.1 and 7.1. Section 8, finally, illustrates the results by means of a numerical simulation, and Section 9 draws conclusions.

Notation: $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$ denote the reals, nonnegative reals, and positive reals, respectively. $\mathbb{N}$ denotes the natural numbers and $\mathbb{N}_0$ the naturals including 0. If $\alpha \in \mathbb{R}$, then $|\alpha|$ is its absolute value. For a set A, its image under f is denoted by $f(A) = \{f(a) : a \in A\}$. One-sided limits of a function f at time instant T from above or below are written as $\lim_{t \to T^+} f(t)$ or $\lim_{t \to T^-} f(t)$, respectively. 'Almost everywhere' is abbreviated as 'a.e.'.

2. Preliminaries and problem statement

This section introduces the considered problem of signal differentiation, and provides formal definitions for important properties of differentiators.

2.1. Signal differentiation

Consider the problem of computing the derivative of a function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ based on a noisy measurement $u = f + \eta$, under the assumption that uniform bounds N and L for the noise η and for the second derivative $\ddot{f}$, respectively, exist. More precisely, let $\mathcal{F}$ denote the set of functions $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ such that f is differentiable and $\dot{f}$ is Lipschitz continuous on $\mathbb{R}_{\geq 0}$. Therefore, if $f \in \mathcal{F}$ then the second derivative $\ddot{f}$ exists almost everywhere (a.e.) due to Rademacher's Theorem. The corresponding classes of signals to consider, from which the measurements are generated, are hence given by

$$\mathcal{F}_L = \{f \in \mathcal{F} : |\ddot{f}(t)| \leq L \text{ a.e. on } \mathbb{R}_{\geq 0}\} \tag{1a}$$

$$\mathcal{E}_N = \{\eta \text{ measurable} : |\eta(t)| \leq N \text{ for all } t \geq 0\}. \tag{1b}$$

Write $\mathcal{F}_L + \mathcal{E}_N = \{f + \eta : f \in \mathcal{F}_L, \eta \in \mathcal{E}_N\}$ for the set of inputs u with fixed L and N. All possible inputs for the differentiator then belong to the set

$$\mathcal{U} = \bigcup_{\substack{L \geq 0 \\ N \geq 0}} (\mathcal{F}_L + \mathcal{E}_N). \tag{2}$$

Note that $\mathcal{U}$ includes all bounded measurable functions.

A differentiator is an operator $\mathcal{D} : \mathcal{U} \to (\mathbb{R}_{\geq 0} \to \mathbb{R})$ that maps the measured signal u to an estimate $\mathcal{D}u$ for the derivative of f. This estimate is hereafter denoted by $y = \mathcal{D}u$.

Definition 2.1. The differentiator $\mathcal{D}$ is said

- to be causal, if $[\mathcal{D}u_1](T) = [\mathcal{D}u_2](T)$ whenever $u_1(t) = u_2(t)$ for all $t \in [0, T]$;
- to be linear, if $\mathcal{D}(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 \mathcal{D}u_1 + \alpha_2 \mathcal{D}u_2$ holds for $u_1, u_2 \in \mathcal{U}$ and $\alpha_1, \alpha_2 \in \mathbb{R}$.

Two further important properties are exactness and robustness, as introduced by Levant (1998). To formally define these properties, the worst-case differentiation error is introduced next.

2.2. Worst-case differentiation error

For future reference, for every $R \geq 0$, define the class of signals with bounded second derivative that, in addition, have a bounded initial value and initial derivative, as follows

$$\mathcal{F}_L^R := \{ f \in \mathcal{F}_L : |f(0)| \leq R, |\dot{f}(0)| \leq R \}. \tag{3}$$

Definition 2.2. Let $L, N \in \mathbb{R}_{\geq 0}$. A differentiator $\mathcal{D}$ is said to have

- worst-case error $M_N^f(t)$ from time $t \geq 0$ for a signal $f \in \mathcal{F}_L$ with noise bound N if

$$M_N^f(t) = \sup_{\substack{u = f + \eta \\ \eta \in \mathcal{E}_N}} \sup_{\tau \geq t} |\dot{f}(\tau) - [\mathcal{D}u](\tau)|; \tag{4}$$

- worst-case error $M_N^{L,R}(t)$ from time $t \geq 0$ over the signal class $\mathcal{F}_L^R$ with noise bound N if

$$M_N^{L,R}(t) = \sup_{f \in \mathcal{F}_L^R} M_N^f(t). \tag{5}$$

Clearly, $M_N^f(t)$ and hence also $M_N^{L,R}(t)$ are non-increasing with respect to t and non-decreasing with respect to N. The latter is also non-decreasing with respect to L and R.

The worst-case error $M_N^{L,R}(t)$ for a fixed time instant t may be an unbounded function of R. This can be seen, e.g., by considering a linear differentiator and is the reason for restricting considerations to the signal class $\mathcal{F}_L^R \subset \mathcal{F}_L$ in Definition 2.2. Using the worst-case differentiation error from time t, different notions of exactness may now be defined.

2.3. Exactness

Exactness, as introduced by Levant (1998), is an important property of differentiators. It guarantees that noise-free signals in a certain class, specifically $\mathcal{F}_L$ in the following, are differentiated exactly, i.e., that the worst-case differentiation error for the noise-free case, $M_0^{L,R}(t)$, is zero for certain values of t. For a given differentiator, this property may sometimes be established only after a certain transient period. Hence, distinctions between different convergence behaviors are made.

Definition 2.3. A differentiator $\mathcal{D}$ is said to be

- exact in finite time over $\mathcal{F}_L$ if for every $R \geq 0$ there exists a t_R such that $M_0^{L,R}(t_R) = 0$;
- exact in fixed time over $\mathcal{F}_L$ if there exists t such that $M_0^{L,R}(t) = 0$ for all $R \geq 0$;
- exact from the beginning over $\mathcal{F}_L$ if $M_0^{L,R}(t) = 0$ for all $R \geq 0$ and all $t > 0$;
- not exact over $\mathcal{F}_L$ if it is not exact in finite time over $\mathcal{F}_L$.

Note that the time instant $t = 0$ is excluded when speaking about exactness from the beginning, since a causal differentiator only knows $f(0)$ at this point. This makes it impossible to deduce $\dot{f}(0)$; specifically, $M_0^{L,R}(0) \geq R$ for all causal differentiators.

For illustration purposes, consider the well-known robust exact differentiator (RED) proposed by (Levant, 1998)

$$\dot{y}_1 = \lambda_1 L^{\frac{1}{2}} |u - y_1|^{\frac{1}{2}} \operatorname{sign}(u - y_1) + y_2 \qquad y_1(0) = u(0)$$

$$\dot{y}_2 = \lambda_2 L \operatorname{sign}(u - y_1) \qquad y_2(0) = 0 \tag{6}$$

with input u, output $y = y_2$, and constant positive parameters λ_1 and $\lambda_2 > 1$. With its solutions understood in the sense of Filippov (1988), this differentiator is exact in finite time over $\mathcal{F}_L$ for sufficiently large λ_1. In particular, for $\lambda_1 \geq \sqrt{8\lambda_2}$, one has $t_R = \frac{R}{(\lambda_2 - 1)L}$, cf. Seeber et al. (2018, Theorem 5).

While the RED is exact in finite time, the uniform robust exact differentiator (URED) proposed by Cruz-Zavala et al. (2011) is exact in fixed time. In contrast, the linear high gain differentiator (HGD) in state–space form as considered by Vasiljevic and Khalil (2008) is not exact. The only existing differentiator that is exact from the beginning is the Euler differentiator $\mathcal{D}u = \dot{u}$. It is well-defined only in the absence of noise, however, i.e., only for $u \in \mathcal{F}_L$. In Section 5, an original differentiator is constructed that is exact from the beginning and is well-defined for input signals in $\mathcal{U}$.

2.4. Robustness

A concept that is closely connected with exactness is robustness. In Levant (1998), a differentiator is said to be robust if the differentiator output $\mathcal{D}(f + \eta)$ tends to $\mathcal{D}f$ uniformly as the uniform bound N on the noise tends to zero. To define robustness here, the following quantities (almost) analogous to $M_N^f(t)$ and $M_0^{L,R}(t)$ are introduced to quantify the deviation between $\mathcal{D}u$ and $\mathcal{D}f$:

$$Q_N^f(t) = \sup_{\substack{u = f + \eta \\ \eta \in \mathcal{E}_N}} \sup_{\tau \geq t} |[\mathcal{D}f](\tau) - [\mathcal{D}u](\tau)|, \tag{7}$$

$$Q^{L,R}(t) = \limsup_{N \to 0^+} \sup_{f \in \mathcal{F}_L^R} Q_N^f(t). \tag{8}$$

With these abbreviations, robustness may then be defined in a similar style as exactness.

**Definition 2.4.** A differentiator $\mathcal{D}$ is said to be

- robust in finite time over $\mathcal{F}_L$, if for every $R \geq 0$ there exists a t_R such that $Q^{L,R}(t_R) = 0$;
- robust in fixed time over $\mathcal{F}_L$, if there exists t such that $Q^{L,R}(t) = 0$ for all $R \geq 0$;
- robust almost from the beginning over $\mathcal{F}_L$, if $Q^{L,R}(t) = 0$ for all $R \geq 0$ and all $t > 0$;
- robust from the beginning over $\mathcal{F}_L$, if $Q^{L,R}(0) = 0$ for all $R \geq 0$;
- not robust over $\mathcal{F}_L$ if it is not robust in finite time over $\mathcal{F}_L$.

Unlike for exactness, the time instant $t = 0$ is not excluded when speaking about robustness *from the beginning*. However, it is impossible for a causal differentiator to be both robust and exact from the beginning as shown in the following, which motivates the additional notion of robustness *almost from the beginning*. All differentiators previously mentioned, i.e., the HGD, the RED, and the URED, are robust from the beginning.

2.5. Exactness and robustness

Some basic relationships between exactness and robustness are next established.

Proposition 2.5. *Let $L \geq 0$ and let $\mathcal{D}$ be a causal differentiator. Then, $\mathcal{D}$ is either not robust from the beginning or not exact from the beginning over $\mathcal{F}_L$.*
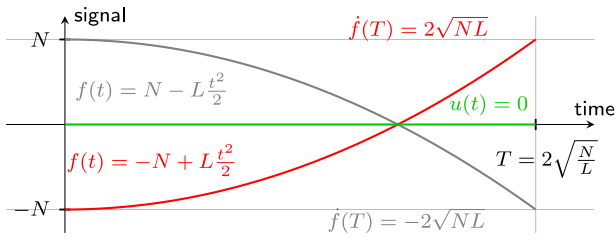
Fig. 1. Input signal limiting the worst-case error of any causal differentiator: for zero input $u(t)$, no differentiator can distinguish between the two illustrated signals $f(t)$, leading to a worst-case error of at least $|\dot{f}(T)| = 2\sqrt{NL}$ at time instant T.

Fig. 2. Input signal limiting the worst-case error of any causal and exact differentiator: for differentiable input $u(t) = -f(t)$, a causal exact differentiator makes an error of at least $2|\dot{f}(T)| = 2\sqrt{2NL}$ at time instant T.

Proof. Suppose to the contrary that $M_0^{L,1}(t) = 0$ for all $t > 0$ and that $Q^{L,0}(0) = 0$. Consider the function $g(t) = t$. Clearly, $g \in \mathcal{F}_L^1$, and hence $[\mathcal{D}g](t) = 1$ for all $t > 0$ due to exactness. Let $N > 0$, and choose the signal $f = 0 \in \mathcal{F}_L^0$ and noise $\eta(t) = \min\{N, t\}$. This yields $u(t) = f(t) + \eta(t) = g(t)$ for $t \in [0, N]$, and hence $[\mathcal{D}u](N) = 1$. Consequently, $Q_N^f(0) \geq Q_N^f(N) \geq 1$, for all $N > 0$, which yields the contradiction $Q^{L,0}(0) \geq 1$. ■

The next proposition shows the well-known fact (cf. Levant, 1998) that linear differentiators cannot be both robust and exact over $\mathcal{F}_L$ for $L > 0$.

Proposition 2.6. *Let $L > 0$ and let $\mathcal{D}$ be a linear differentiator. Then, $\mathcal{D}$ is either not robust or not exact over $\mathcal{F}_L$.*

Proof. Suppose to the contrary that $\mathcal{D}$ is robust and exact over $\mathcal{F}_L$ in finite time. Let $T \geq 1$ be such that $M_0^{L,0}(T) = 0$ and $Q^{L,0}(T) = 0$. Let $\omega_k = 2\pi k$ for $k = 1, 2, \ldots$ and consider the sequence of noise inputs $\eta_k(t) = \min\{1, t^2\}\sin(\omega_k t)/\omega_k$. One verifies that $\eta_k(0) = \dot{\eta}_k(0) = 0$. Since $L > 0$, there furthermore exists an $\alpha_k > 0$ for each η_k such that $\alpha_k \eta_k \in \mathcal{F}_L^0$, since $\ddot{\eta}_k$ is uniformly bounded. Due to linearity, $\mathcal{D}$ then differentiates also η_k exactly for $t \geq T$. Thus, for $f = 0$ and all k,

$$Q_{1/\omega_k}^f(T) \geq \sup_{\tau \geq T} |[\mathcal{D}\eta_k](\tau)| = \sup_{\tau \geq T} |\dot{\eta}_k(\tau)| = 1. \tag{9}$$

Since $1/\omega_k \to 0$ as $k \to \infty$, this contradicts $Q^{L,0}(T) = 0$. ■

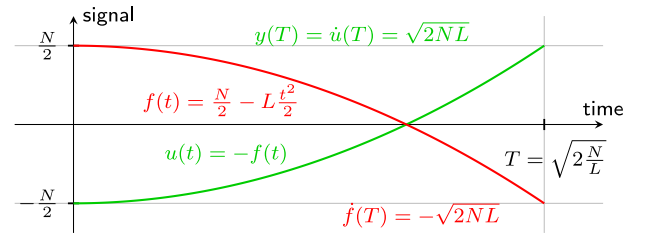## 3. Performance limits and differentiation accuracy

To motivate the notion of optimal differentiation accuracy, worst-case error lower bounds of causal and causal exact differentiators are shown next.

3.1. Worst-case error lower bound of causal differentiators

The main limitation of the worst-case differentiation error of *any* causal differentiator is shown in the following proposition. It stems from the fact that the differentiator, when fed zero input $u = 0$, cannot distinguish between the zero function and a parabola arc $Lt^2/2 - N$ staying within the noise bound. In Section 4, a differentiator is constructed that attains this lower bound.

Proposition 3.1. *Let $\mathcal{D}$ be a causal differentiator and let $L, N, R \geq 0$. Then, $M_N^{L,R}(\tau) \geq 2\sqrt{NL}$ for all $\tau \geq 0$.*

Remark 3.2. The rationale of the proof is sketched in Fig. 1. ◦

Proof. If $L = 0$ or $N = 0$, then there is nothing to prove. Otherwise, define $h_\kappa : [0, 2\kappa] \to [0, L\kappa^2]$ as

$$h_\kappa(t) = \begin{cases} L\dfrac{t^2}{2} & t \in [0, \kappa) \\ L\kappa^2 - L\dfrac{(t-2\kappa)^2}{2} & t \in [\kappa, 2\kappa] \end{cases} \tag{10}$$

and let $\kappa = \sqrt{N/L}$. Then, $h_\kappa(0) = \dot{h}_\kappa(0) = \dot{h}_\kappa(2\kappa) = 0$ and $h_\kappa(2\kappa) = N$. Let $T = \tau + 4\kappa$ and consider the functions

$$g_1(t) = \begin{cases} 0 & t \in [0, \tau) \\ -h_\kappa(t - \tau) & t \in [\tau, \tau + 2\kappa) \\ -N + L\dfrac{(t-\tau-2\kappa)^2}{2} & t \geq \tau + 2\kappa, \end{cases} \tag{11a}$$

$$g_2(t) = \begin{cases} -g_1(t) & t \leq T \\ -N & \text{otherwise} \end{cases} \tag{11b}$$

It is straightforward to verify that for all $t \geq 0$, $|\ddot{g}_1(t)| \leq L$ a.e. and $|g_2(t)| \leq N$ hold, so that $g_1 \in \mathcal{F}_L^0 \subseteq \mathcal{F}_L^R$ and $g_2 \in \mathcal{E}_N$. Choosing either $f = g_1$ and $\eta = g_2$ or $f = -g_1$ and $\eta = -g_2$ results in identical (zero) input for all $t \leq T$. Since the differentiator is causal, its output $y(T)$ must be the same in either case. We have $\lim_{t \to T^-} 2\dot{g}_1(t) = 4\sqrt{NL}$. Hence, $\max\{M_N^{g_1}(\tau), M_N^{-g_1}(\tau)\} \geq 2\sqrt{NL}$ since $\tau \leq T$ and, as a consequence, $M_N^{L,R}(\tau) \geq 2\sqrt{NL}$. ■

3.2. Worst-case error lower bound of exact differentiators

Exact differentiators have a more restrictive bound on their worst-case differentiation error than the bound in Proposition 3.1. This is due to the fact that also some noise signals are differentiated exactly, as also noted by Levant (1998). Section 5 shows a differentiator that attains this bound.

Proposition 3.3. *Let $\mathcal{D}$ be a causal differentiator and let $L, N, R \geq 0$. If $\mathcal{D}$ is exact in finite time over $\mathcal{F}_L$, then $M_N^{L,R}(\tau) \geq 2\sqrt{2NL}$ for all $\tau \geq 0$.*

Remark 3.4. The rationale of the proof is sketched in Fig. 2. ◦

Remark 3.5. A similar bound is shown in Levant et al. (2017, Proposition 1) for arbitrary differentiation orders. However, that bound is also valid for acausal differentiators, and hence the lower bound $2\sqrt{NL}$ is obtained there for a first-order differentiator. A tighter bound based on Levant (1998) is mentioned by Fraguela et al. (2012) in a footnote. ◦

Proof. For $N = 0$ or $L = 0$, the statement is trivial. Consider hence arbitrary $L, N > 0$, let t_0 be such that $M_0^{L,0}(t_0) = 0$ from Definition 2.3, and consider any $\tau \geq t_0$. It will be shown that $M_N^{L,0}(\tau) \geq 2\sqrt{2NL}$, from which the claim follows from the fact that $M_N^{L,R}(t)$ is non-decreasing with respect to R and non-increasing with respect to t. Define $h_\kappa : [0, 2\kappa] \to [0, L\kappa^2]$ as

in (10) with $\kappa = \sqrt{N/L}$. Let $T = \tau + (2 + \sqrt{2})\kappa$ and consider the functions

$$g_1(t) = \begin{cases} 0 & t \in [0, \tau) \\ -\dfrac{h_\kappa(t-\tau)}{2} & t \in [\tau, \tau + 2\kappa) \\ -\dfrac{N}{2} + L\dfrac{(t-\tau-2\kappa)^2}{2} & t \geq \tau + 2\kappa, \end{cases} \tag{12a}$$

$$g_2(t) = \begin{cases} 2g_1(t) & t \leq T \\ N & \text{otherwise} \end{cases} \tag{12b}$$

It is straightforward to verify that $|\ddot{g}_1(t)| \leq L$ a.e. for $t \geq 0$, that $g_1 \in \mathcal{F}_L^0$, and that $g_2 \in \mathcal{E}_N$. From exactness, $[\mathcal{D}g_1](t) = \dot{g}_1(t)$ for all $t \geq \tau$. Choosing $f = -g_1$, $\eta = g_2$ yields $u(t) = f(t) + \eta(t) = g_1(t)$ for $t \in [0, T]$; hence, also $[\mathcal{D}u](t) = \dot{g}_1(t)$ for $t \in [\tau, T]$, since the differentiator is causal. Consequently, $[\mathcal{D}u](T) = \sqrt{2NL}$ whereas $\dot{f}(T) = -\sqrt{2NL}$. This establishes that $M_N^f(\tau) \geq 2\sqrt{2NL}$ with $f \in \mathcal{F}_L^0$. Therefore, $M_N^{L,0}(\tau) \geq 2\sqrt{2NL}$. ∎

3.3. Differentiation accuracy

As one can see from Proposition 3.1, the lower bound of the worst-case differentiation error is proportional to $\sqrt{NL}$. In Vasiljevic and Khalil (2008), the HGD is also shown to achieve a similar proportionality with properly chosen parameters which, however, depend on N.

If the noise bound is unknown, it is desirable that this proportionality be maintained either for all or sufficiently small N. This motivates the following notions of global or asymptotic differentiation accuracy, respectively. The latter notion is loosely consistent with similar notions in the context of sliding mode differentiatiors, where asymptotic accuracy also refers to the asymptotic behavior of the differentiation error with respect to (small) noise bounds N.

Definition 3.6. A differentiator $\mathcal{D}$ is said to have

(a) asymptotic accuracy $C_L \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ for signals in $\mathcal{F}_L$, if C_L is the infimum of all numbers C with the property that there exist $\epsilon > 0$ and a function $\mathcal{T} : \mathbb{R}_{\geq 0} \times [0, \epsilon] \to \mathbb{R}_{\geq 0}$ continuous in its second argument such that

$$M_N^{L,R}[\mathcal{T}(R, N)] \leq C\sqrt{NL} \tag{13}$$

for all $N \in [0, \epsilon]$ and $R \geq 0$;

(b) global accuracy $\overline{C}_L \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ for signals in $\mathcal{F}_L$, if $\overline{C}_L$ is the infimum of all numbers C with the property that there exists a function $\mathcal{T} : \mathbb{R}_{\geq 0}^2 \to \mathbb{R}_{\geq 0}$ continuous in its second argument such that (13) holds for all $R, N \geq 0$.

Note that $\mathcal{T}$ in this definition can be considered as a kind of convergence-time function. For given initial value bound R and noise bound N, it yields the time $\mathcal{T}(R, N)$ after which the worst-case differentiation error is bounded by $C\sqrt{NL}$. Asymptotic accuracy is a property pertaining to small values of the noise amplitude, whereas global accuracy applies to all noise amplitudes and is suitable when N is unknown. Clearly, $\overline{C}_L \geq C_L$ holds by definition, and it is also possible that C_L is finite when $\overline{C}_L$ is infinite. Important differentiators with a finite asymptotic accuracy are the RED and the URED. The former also has finite global accuracy due to its homogeneity properties, while the latter does not.

The connection between robustness, exactness and accuracy is established by the following proposition, which shows that finite asymptotic accuracy implies robustness and exactness.

Proposition 3.7. Let $L \geq 0$ and let $\mathcal{D}$ be a differentiator with asymptotic accuracy $C_L \in \mathbb{R}_{\geq 0}$. Then, $\mathcal{D}$ is both robust and exact in finite time over $\mathcal{F}_L$.

Remark 3.8. From (13), one can see that $t_R = \mathcal{T}(R, 0)$ in Definition 2.3 with any $\mathcal{T}$ as in Definition 3.6. ∘

Proof. In order to prove robustness in finite time, note that $Q_N^f(t) \leq M_N^f(t) + M_0^f(t)$ for all t and $f \in \mathcal{F}_L$. Let $R \geq 0$, $\epsilon > 0$, and let $\mathcal{T}$ be the function from Definition 3.6 for some $C > C_L$. Consider the image $\mathcal{I} = \mathcal{T}(R, [0, \epsilon])$, which is a compact interval due to continuity of $\mathcal{T}$ with respect to its second argument, and set $t_R = \max \mathcal{I}$. For all $N \in [0, \epsilon]$, then $M_N^{L,R}(t_R) \leq C\sqrt{NL}$ since $M_N^{L,R}(\cdot)$ is non-increasing with respect to its argument. Hence,

$$\sup_{f \in \mathcal{F}_L^R} Q_N^f(t_R) \leq M_N^{L,R}(t_R) + M_0^{L,R}(t_R) \leq C\sqrt{NL}, \tag{14}$$

holds for all $N \in [0, \epsilon]$, proving that $Q^{L,R}(t_R) = 0$. Exactness follows by noting that also $M_0^{L,R}(t_R) = 0$. ∎

Note that, although only robustness in finite time may be concluded from finite asymptotic accuracy, most existing differentiators in fact exhibit robustness from the beginning. In contrast, exactness from the beginning, as pointed out before, is a very strong property which is not known to be obtained by any existing robust differentiators up to now.

An immediate consequence of the previous proposition along with Proposition 2.6 is the fact that linear differentiators cannot have finite asymptotic accuracy.

Proposition 3.9. Let $L > 0$ and let $\mathcal{D}$ be a linear differentiator. Then, $\overline{C}_L = C_L = \infty$.

Proof. Since $\overline{C}_L \geq C_L$, assume to the contrary that C_L is finite. Then, $\mathcal{D}$ is robust and exact in finite time over $\mathcal{F}_L$ by Proposition 3.7, contradicting Proposition 2.6. ∎

From Proposition 3.1 it is clear that $C_L \geq 2$ bounds the asymptotic accuracy of any causal differentiator from below. This bound is too conservative, however; a tighter lower bound, which will be shown to be achievable in Section 5, is obtained from Propositions 3.7 and 3.3.

Proposition 3.10. Let $L > 0$ and let $\mathcal{D}$ be a causal differentiator. Then, its global and asymptotic accuracy satisfy $\overline{C}_L \geq C_L \geq 2\sqrt{2}$.

Proof. It is sufficient to show $C_L \geq 2\sqrt{2}$. To that end, assume to the contrary that $C_L < 2\sqrt{2}$. Then, C_L is finite and $\mathcal{D}$ is exact in finite time according to Proposition 3.7. From Proposition 3.3, then $M_N^{L,R}(t) \geq 2\sqrt{2NL}$ for all $N > 0$, yielding the contradiction $C_L \geq 2\sqrt{2}$. ∎

For $L = 0$, finally, it is impossible to achieve finite accuracy.

Proposition 3.11. Let $\mathcal{D}$ be a causal differentiator. Then, its global and asymptotic accuracy satisfy $\overline{C}_0 = C_0 = \infty$.

Proof. Suppose to the contrary that $C_L < \infty$ for $L = 0$ and let $R > 0$. From (13), then $N, \tau \in \mathbb{R}_{>0}$ exist such that $M_N^{0,R}(\tau) = 0$. Let $\delta = \min\{N/\tau, R\}$ and consider $g(t) = \delta t$. Clearly, $g \in \mathcal{F}_0^R$, and hence $[\mathcal{D}g](\tau) = \dot{g}(\tau) = \delta$. Now choose $f = 0 \in \mathcal{F}_0^R$ and $\eta(t) = \min\{g(t), N\}$ to obtain $u(t) = f(t) + \eta(t) = g(t)$ for $t \leq \tau$, leading to the contradiction $[\mathcal{D}u](\tau) - \dot{f}(\tau) = \delta$ due to causality, i.e., $M_N^{0,R}(\tau) \geq \delta > 0$. ∎

4. Linear differentiators with best worst-case error

As a preliminary result for ultimately constructing a robust exact differentiator with optimal differentiation accuracy, this

section shows how a linear differentiator with lowest possible worst-case error $2\sqrt{NL}$, as given by Proposition 3.1, can be constructed.

Consider the following differentiator $\mathcal{D}$ with a positive parameter T and output $y = \mathcal{D}u$ given by

$$y(t) = \begin{cases} \frac{u(t) - u(t-T)}{T} & \text{if } t \geq T \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

If the noise bound N is precisely known, then this differentiator achieves the best possible worst-case differentiation error by suitable selection of T.

Theorem 4.1. *Let $N > 0$, $L > 0$ and consider the differentiator (15) with $T = 2\sqrt{N/L}$. Then, $M_N^{L,R}(\tau) = 2\sqrt{NL}$ for all $R \geq 0$ and all $\tau \geq T$.* △

The proof uses the following two lemmata.

Lemma 4.2. *Let $L \in \mathbb{R}_{\geq 0}$ and $f \in \mathcal{F}_L$. Then,*

$$\left| f(t - \sigma) - f(t) + \dot{f}(t)\sigma \right| \leq \frac{L\sigma^2}{2} \tag{16}$$

holds for all $t \geq 0$ and all $\sigma \in [0, t]$.

Proof. Consider an arbitrary $t \geq 0$ and define the function $g(\sigma) = f(t - \sigma) - f(t) + \dot{f}(t)\sigma$. Clearly, g is a.e. twice differentiable, its second derivative satisfies $\ddot{g}(\sigma) = \ddot{f}(t - \sigma)$ a.e. in $[0, t]$, and $g(0) = \dot{g}(0) = 0$. The lemma's claim $|g(\sigma)| \leq L\sigma^2/2$ follows by double integration of $\ddot{g}$ using the bound $|\ddot{g}(\sigma)| \leq L$ starting from $\sigma = 0$. ■

Lemma 4.3. *Let $T > 0$ and consider the differentiator $\mathcal{D}$ with output $y = \mathcal{D}u$ defined in (15). Suppose that $f \in \mathcal{F}_L$ and $\eta \in \mathcal{E}_N$. Then, $|y(t) - \dot{f}(t)| \leq \frac{2N}{T} + \frac{LT}{2}$ holds for all $t \geq T$.*

Remark 4.4. As will be shown later, the bound is also valid when the parameter T is time varying. This motivates the adaptation of that parameter to eventually obtain a robust exact differentiator in Section 5. A related approach is proposed by Levant (2007), where finite differences with adaptation of the sampling time are used in a closed-loop sliding mode control scheme. ○

Proof. From (15), $|u(t - T) - u(t) + y(t)T| = 0$ for $t \geq T$, and hence $|f(t - T) - f(t) + y(t)T| \leq 2N$, since $|u(\tau) - f(\tau)| \leq N$ for all τ. Setting $\sigma = T$ in (16) in Lemma 4.2 and combining the two inequalities yields $\left| y(t)T - \dot{f}(t)T \right| \leq LT^2/2 + 2N$. The claim follows after dividing by T. ■

Using Lemma 4.3, the optimal worst-case accuracy of the linear differentiator (15) may now be shown.

PROOF of Theorem 4.1. Set $T = 2\sqrt{N/L}$ in Lemma 4.3 to obtain

$$M_N^{L,R}(t) \leq \frac{2N}{T} + \frac{LT}{2} = 2\sqrt{NL} \tag{17}$$

for $t \geq T$. Equality is concluded by noting that $M_N^{L,R}(t) \geq 2\sqrt{NL}$ due to Proposition 3.1. ■

5. Robust exact differentiators with optimal accuracy

The problem with differentiator (15) is that its tuning requires knowledge of the noise amplitude. If the noise affecting its input is actually of lower amplitude than the design parameter N used in the differentiator's construction, achieving an optimal worst-case differentiation error cannot be ensured. The main idea for

obtaining an optimal differentiator based on knowledge of only an upper bound on the noise amplitude or directly without any knowledge on the noise amplitude is to obtain a reasonable estimate $\hat{N}$ for this amplitude.

5.1. Proposed differentiator

From the above considerations, a robust exact differentiator with best possible worst-case accuracy is constructed as the linear time-varying differentiator[1]

$$y(t) = \begin{cases} 0 & \text{if } t = 0 \\ \lim_{T \to 0^+} \dfrac{u(t) - u(t-T)}{T} & \text{if } t > 0,\, \hat{T}(t) = 0 \\ \dfrac{u(t) - u(t - \hat{T}(t))}{\hat{T}(t)} & \text{if } t > 0,\, \hat{T}(t) > 0 \end{cases} \tag{18a}$$

with an adaptation of the time difference $\hat{T}(t)$ according to

$$\hat{T}(t) = \min\left\{ t, \overline{T}, 2\gamma(t)\sqrt{\frac{\hat{N}(t)}{L}} \right\} \tag{18b}$$

wherein the parameter $\overline{T} \in \mathbb{R}_{>0} \cup \{\infty\}$ constrains the time difference from above, $\gamma : \mathbb{R}_{\geq 0} \to [1, \overline{\gamma}]$ is an arbitrary function constrained by a constant parameter $\overline{\gamma} \geq 1$, i.e.,

$$\gamma(t) \in [1, \overline{\gamma}] \qquad \text{for all } t, \tag{18c}$$

and $\hat{N}(t)$ is an estimate for the noise amplitude determined from the measurement u according to

$$\hat{N}(t) = \frac{1}{2} \sup_{\substack{T \in (0, \overline{T}] \\ T \leq t \\ \sigma \in [0, T]}} \left(|Q(t, T, \sigma)| - \frac{L\sigma(T - \sigma)}{2} \right) \tag{18d}$$

with the abbreviation $Q(t, T, \sigma)$ defined as

$$Q(t, T, \sigma) = u(t - \sigma) - u(t) + \frac{u(t) - u(t - T)}{T}\sigma. \tag{18e}$$

Note that $\hat{N}(t) \geq 0$ since $Q(t, T, 0) = 0$ and hence the argument of the supremum is zero for $\sigma = 0$.

The parameter $\overline{T}$ may be considered to be a window-length parameter, since computing $y(t)$ requires evaluation of u only on the interval $[t - \overline{T}, t]$. If a finite value is chosen for this parameter, it allows to limit the proposed differentiator's computational complexity.

The function γ satisfying $1 \leq \gamma(t) \leq \overline{\gamma}$ for all t is a degree of freedom parametrizing a whole family of optimal robust exact differentiators. In the presence of sampled measurements, in Section 7.1, this degree of freedom will be exploited to choose $\hat{T}(t)$ as a multiple of the sampling time, yielding a straightforward discrete-time implementation.

The next main theorem, which is proven in Section 5.5, establishes that (i) the output $y(t)$ of this differentiator is always well-defined and, in particular, the one-sided limit occurring in (18a) exists whenever $\hat{T}(t) = 0$ and $t > 0$, (ii) this differentiator is exact from the beginning and robust almost from the beginning with any $\overline{\gamma} \geq 1$, and (iii) it has optimal asymptotic accuracy $C_L = 2\sqrt{2}$ for $\overline{\gamma} \in [1, 1 + \sqrt{2}]$ and also optimal global accuracy if moreover $\overline{T} = \infty$.

Theorem 5.1. *Let $L \in \mathbb{R}_{>0}$ and consider the differentiator $\mathcal{D}$ with output $y = \mathcal{D}u$ defined by (18) with parameters $\overline{\gamma} \geq 1$ $\overline{T} \in \mathbb{R}_{>0} \cup \{\infty\}$ and an arbitrary function $\gamma : \mathbb{R}_{\geq 0} \to [1, \overline{\gamma}]$. Then, the following statements are true:*

---

[1] See Section 7 for the practical, sample-based implementation of the proposed continuous-time differentiator.

(a) *The output $y = \mathcal{D}u$ is well-defined for all $u \in \mathcal{U}$.*
(b) *$\mathcal{D}$ is robust almost from the beginning and exact from the beginning over $\mathcal{F}_L$.*
(c) *If $\overline{\gamma} \in [1, 1 + \sqrt{2}]$, then $\mathcal{D}$ achieves optimal asymptotic accuracy $C_L = 2\sqrt{2}$; specifically, $M_N^{L,R}(t) \leq 2\sqrt{2NL}$ holds for all $N \in [0, L\overline{T}^2/2)$ and all $t > \sqrt{2N/L}$.*
(d) *If $\overline{\gamma} \in [1, 1 + \sqrt{2}]$ and $\overline{T} = \infty$, then $\mathcal{D}$ achieves optimal global accuracy $\overline{C}_L = 2\sqrt{2}$.* △

Remark 5.2 (*Tuning*). If a (crude) upper bound $\overline{N}$ for the noise amplitude is known, then a robust exact differentiator with bounded complexity that is optimal for all $N < \overline{N}$ may be obtained by choosing the parameters $\overline{\gamma} = 1$ and $\overline{T} = \sqrt{2\overline{N}/L}$. ○

Remark 5.3. The proposed differentiator achieves the best possible exactness and robustness features, given that it is impossible to achieve both properties from the beginning, according to Proposition 2.6, as well as the best possible accuracy, according to Proposition 3.10. ○

The proof of Theorem 5.1 requires the analysis of several properties, as performed in the next subsections. The proof itself is given afterwards, in Section 5.5.

5.2. Adaptation of the time-difference parameter

To motivate the adaptation of $\hat{T}(t)$ in (18b), the following lemma is obtained essentially as a corollary of Lemma 4.3.

Lemma 5.4. *Let $N, L \in \mathbb{R}_{\geq 0}$ and consider the differentiator $\mathcal{D}$ with output $y = \mathcal{D}u$ defined in (18). Suppose that $f \in \mathcal{F}_L$, $\eta \in \mathcal{E}_N$, and $u = f + \eta$. Then,*

$$\left| y(t) - \dot{f}(t) \right| \leq \frac{2N}{\hat{T}(t)} + \frac{L\hat{T}(t)}{2} \tag{19}$$

holds whenever $t \geq \hat{T}(t) > 0$.

Proof. Replace T by $\hat{T}(t)$ in the proof of Lemma 4.3. ∎

It is straightforward to verify that this lemma yields the optimal error bound $|y(t) - \dot{f}(t)| \leq 2\sqrt{2NL}$ if $\hat{T}(t) = 2\gamma(t)\sqrt{N/L}$ with $\gamma(t) \in [\sqrt{2} - 1, 1 + \sqrt{2}]$ for all t. This fact motivates the structure of $\hat{T}(t)$ in (18b).

Since N is not available, an estimate $\hat{N}(t)$, given in (18d), is used instead of N in the actual computation of $\hat{T}(t)$. The properties of this noise amplitude estimate are analyzed next. It will be shown that it is not necessarily equal to, but always bounded from above by N; this fact intuitively explains the restriction of $\gamma(t)$ to the interval $[1, \overline{\gamma}]$ with $\overline{\gamma} \leq 1 + \sqrt{2}$ as opposed to also allowing values $\gamma(t) < 1$ as above.

5.3. Properties of the noise estimate

This section discusses upper and lower bounds on the noise estimate (18d) and shows its relation to differentiability and growth bound of the measurements.

5.3.1. Upper and lower bounds for $\hat{N}$
To illuminate the noise amplitude estimation in (18d)–(18e) and to obtain an upper bound for $\hat{N}$, the following lemma shows how Q in (18e) is related to the noise bound N.

Lemma 5.5. *Let $L, N \in \mathbb{R}_{\geq 0}$. For any $t \in \mathbb{R}_{>0}$, consider $Q(t, T, \sigma)$ as defined in (18e) with $u \in \mathcal{F}_L + \mathcal{E}_N$. Then,*

$$|Q(t, T, \sigma)| \leq 2N + \frac{L\sigma(T - \sigma)}{2} \tag{20}$$

holds for all $T \in (0, t)$ and all $\sigma \in [0, T]$.

Proof. For arbitrary, fixed $t > 0$ and any function $w \in \{u, f, \eta\}$, define

$$a_T^w(\sigma) = w(t - \sigma) - w(t) + \frac{w(t) - w(t - T)}{T}\sigma. \tag{21}$$

Then, $Q(t, T, \sigma) = a_T^u(\sigma) = a_T^\eta(\sigma) + a_T^f(\sigma)$. An upper bound for a_T^η is given by

$$\begin{aligned}
\left| a_T^\eta(\sigma) \right| &\leq |\eta(t - \sigma)| + \left(1 - \frac{\sigma}{T}\right)|\eta(t)| + \frac{\sigma}{T}|\eta(t - T)| \\
&\leq N + \left(1 - \frac{\sigma}{T}\right)N + \frac{\sigma}{T}N = 2N.
\end{aligned} \tag{22}$$

Moreover, a_T^f is continuously differentiable and satisfies

$$a_T^f(0) = a_T^f(T) = 0, \qquad \ddot{a}_T^f(\sigma) = \ddot{f}(t - \sigma), \tag{23}$$

and hence a.e. $|\ddot{a}_T^f(\sigma)| \leq L$. Subject to these constraints, the extremal is given by

$$\overline{a}_T^f(\sigma) = \frac{L\sigma(T - \sigma)}{2} \tag{24}$$

i.e., $|a_T^f(\sigma)| \leq \overline{a}_T^f(\sigma)$. The proof is concluded by noting that $|a_T^u(\sigma)| \leq |a_T^\eta(\sigma)| + |a_T^f(\sigma)| \leq 2N + \overline{a}_T^f(\sigma)$. ∎

Lemma 5.5 suggests that an estimate for N may be obtained by subtracting the parabola arc $L\sigma(T - \sigma)/2$ from $|Q(t, T, \sigma)|$. Indeed, the proposed estimate $\hat{N}(t)$ in (18d) can be seen to be constructed by taking the supremum of this difference over all $0 < \sigma \leq T \leq \overline{T}$.

Clearly $\hat{N}(t)$ is bounded by N from above by construction; moreover, discontinuities in the noise impose a lower bound. These properties are summarized in the following lemma.

Proposition 5.6. *Let $L, N \in \mathbb{R}_{\geq 0}$, $\overline{T} \in \mathbb{R}_{>0} \cup \{\infty\}$, suppose $u = f + \eta$ with $f \in \mathcal{F}_L$ and $\eta \in \mathcal{E}_N$, and consider a fixed $t \in \mathbb{R}_{>0}$. Define the (right-sided) discontinuity of η at t_0 as*

$$D(t_0) = \limsup_{\tau \to t_0^+} |\eta(\tau) - \eta(t_0)|. \tag{25}$$

Then, $\hat{N}(t)$ as defined in (18d) satisfies

$$\frac{D(t_0)}{2} \leq \hat{N}(t) \leq N \tag{26}$$

for all $t_0 \in [0, t)$ with $t_0 \geq t - \overline{T}$.

Proof. The inequality $\hat{N}(t) \leq N$ follows from the upper bound on $|Q(t, T, \sigma)|$ from Lemma 5.5. To show also $D(t_0) \leq 2\hat{N}(t)$, fix $t_0 \in [0, t)$ with $t_0 \geq t - \overline{T}$, and let (τ_k) be a sequence with $\tau_k \to t_0$, $\tau_k \in [t_0, t)$, and $|\eta(\tau_k) - \eta(t_0)| \to D(t_0)$. Let $\sigma_k = t - \tau_k \geq 0$, $T = t - t_0 > 0$ and define a_T^w as in (21) for $w \in \{u, f, \eta\}$, allowing to write $Q(t, T, \sigma) = a_T^\eta(\sigma) + a_T^f(\sigma)$. Then, $\sigma_k \to T \leq \overline{T}$ and due to the continuity of f (and thus of a_T^f), one has

$$\lim_{k \to \infty} a_T^f(\sigma_k) = 0 = \lim_{k \to \infty} \frac{L\sigma_k(T - \sigma_k)}{2}. \tag{27}$$

Hence, $2\hat{N}(t)$ is bounded from below by

$$2\hat{N}(t) \geq \lim_{k \to \infty} \left| a_T^\eta(\sigma_k) \right| = \lim_{k \to \infty} |\eta(t - \sigma_k) - \eta(t - T)|$$

$$= \lim_{k \to \infty} |\eta(\tau_k) - \eta(t_0)| = D(t_0), \tag{28}$$

concluding the proof. ∎

Depending on whether $\hat{N}(t)$ at a given time instant $t > 0$ is zero or nonzero, two cases need to be distinguished.

5.3.2. Differentiability and growth bound for $\hat{N} = 0$

If $\hat{N}(t)$ is zero, then the estimation (at time instant t) does not detect the presence of noise. In this case, Q in (18e) satisfies

$$|Q(t, T, \sigma)| \leq \frac{L\sigma(T - \sigma)}{2} \tag{29}$$

for $0 < \sigma \leq T \leq \overline{T}$ by definition of $\hat{N}$ in (18d). In the following, it will be shown that a (one-sided) derivative of the measurement u exists in this case, and that u satisfies a similar growth bound as f in Lemma 4.2.

The following lemma shows that $\hat{N}(t) = 0$ implies left-sided differentiability of the noisy measurement u at t.

Lemma 5.7. *Let $L \in \mathbb{R}_{\geq 0}$, suppose that $u \in \mathcal{U}$, and consider a fixed $t \in \mathbb{R}_{>0}$. Let $\mu \in (0, t]$ and suppose Q as defined in (18e) satisfies (29) for all $T \in (0, \mu]$ and all $\sigma \in [0, T]$. Then, the limit*

$$\beta := \lim_{T \to 0^+} \frac{u(t) - u(t - T)}{T} \tag{30}$$

exists.

Proof. Consider the function $h : (0, \mu] \to \mathbb{R}$ defined by

$$h(\sigma) = \frac{u(t) - u(t - \sigma)}{\sigma}. \tag{31}$$

From (18e), one obtains $Q(t, T, \sigma) = (h(\sigma) - h(T))\sigma$. For given $\epsilon > 0$, consider arbitrary σ and T satisfying the inequalities $\min(\epsilon, \mu) > T > \sigma > 0$. Then,

$$|h(\sigma) - h(T)| = \frac{|Q(t, T, \sigma)|}{\sigma} \leq \frac{L(T - \sigma)}{2} < \frac{L\epsilon}{2} \tag{32}$$

which implies existence of the limit $\beta = \lim_{\sigma \to 0^+} h(\sigma)$. ∎

The next lemma shows that the measurements satisfy a similar growth bound as f in Lemma 4.2.

Lemma 5.8. *Suppose that the conditions of Lemma 5.7 are fulfilled and let β be defined as in (30) in that lemma. Then,*

$$|u(t - \sigma) + \beta\sigma - u(t)| \leq \frac{L\sigma^2}{2} \tag{33}$$

holds for all $\sigma \in [0, \mu]$.

Proof. Consider the function $h : [0, \mu] \to \mathbb{R}$ defined by (31) for $\sigma > 0$ and $h(0) = \beta$. According to Lemma 5.7, this function is continuous at $\sigma = 0$ and by definition fulfills the relation $u(t - \sigma) + \beta\sigma - u(t) = \sigma(h(0) - h(\sigma))$. Using (29), one thus obtains

$$|h(0) - h(\sigma)| = \left| \sum_{i=0}^\infty h(\sigma/2^{i+1}) - h(\sigma/2^i) \right|$$

$$\leq \sum_{i=0}^\infty \frac{\left| Q(t, \sigma/2^{i+1}, \sigma/2^i) \right|}{\sigma/2^i}$$

$$\leq \sum_{i=0}^\infty \frac{L}{2} \left[\sigma/2^i - \sigma/2^{i+1} \right] = \frac{L\sigma}{2} \tag{34}$$

for $\sigma \in [0, \mu]$, yielding the claimed inequality. ∎

5.3.3. Growth bound for $\hat{N} > 0$

If $\hat{N}(t)$ is non-zero, then the measurements allow to distinguish noise, whose magnitude N is at least $\hat{N}$. In this case, it is neither possible nor necessary to compute an exact derivative of u at time instant t.

The following lemma shows that a similar inequality as in Lemma 5.8 may nonetheless be obtained with a suitable value of β.

Lemma 5.9. *Let $L, \hat{N} \in \mathbb{R}_{>0}$, let $u \in \mathcal{U}$, and consider a fixed $t \in \mathbb{R}_{>0}$. Let $\hat{T} \geq 2\sqrt{\hat{N}/L}$, define*

$$\beta := \frac{u(t) - u(t - \hat{T})}{\hat{T}}, \tag{35}$$

and suppose that Q as defined in (18e) satisfies

$$\left| Q(t, \hat{\sigma}, \hat{T}) \right| \leq 2\hat{N} + \frac{L\hat{T}(\hat{\sigma} - \hat{T})}{2} \tag{36}$$

for some $\hat{\sigma} \in [\hat{T}, t]$. Then, (33) holds for $\sigma = \hat{\sigma}$.

Remark 5.10. Note that for $\hat{N} = \hat{N}(t)$ as defined in (18d), condition (36) of this lemma is fulfilled for every $\hat{\sigma} \in [\hat{T}, t]$ satisfying $\hat{\sigma} \leq \overline{T}$. ○

Proof. Consider the function $h : (0, t] \to \mathbb{R}$ defined by (31). Then, $\beta = h(\hat{T})$ and

$$\frac{\left| u(t - \hat{\sigma}) + \beta\hat{\sigma} - u(t) \right|}{\hat{\sigma}} = \left| h(\hat{T}) - h(\hat{\sigma}) \right| \leq \frac{|Q(t, \hat{\sigma}, \hat{T})|}{\hat{T}}$$

$$\leq \frac{2\hat{N}}{\hat{T}} + \frac{L(\hat{\sigma} - \hat{T})}{2} \leq \frac{L\hat{\sigma}}{2} \tag{37}$$

since $\hat{T}^2 \geq 4\hat{N}/L$ implies $2\hat{N}/\hat{T} \leq L\hat{T}/2$. ∎

5.4. Worst-case error upper bound

To prove the optimal accuracy, the following lemma shows how a bound for the differentiation error may be obtained from the growth bounds proven in Lemmata 5.8 and 5.9.

Lemma 5.11. *Let $L \in \mathbb{R}_{>0}$, $N \in \mathbb{R}_{>0}$ and $\beta \in \mathbb{R}$. Define $\ell := \sqrt{2N/L}$, let $t \geq \ell$ and suppose that $u = f + \eta$ with $f \in \mathcal{F}_L$ and $\eta \in \mathcal{E}_N$ satisfies*

$$|u(t - \sigma) - u(t) + \beta\sigma| \leq \frac{L\sigma^2}{2} \tag{38}$$

for $\sigma = \ell + x$ with $x \in [0, \Delta]$ for some $\Delta \in [0, \ell]$. Then, $|\beta - \dot{f}(t)| \leq 2\sqrt{2NL} + L\frac{\Delta}{2}$.

Remark 5.12. The degree of freedom Δ introduced in this lemma is zero in the continuous-time case, which yields the desired optimal bound (cf. Proposition 3.3). Later, Δ will be used to derive an error bound also for a discrete-time implementation of the differentiator. ○

Proof. According to Lemma 4.2, f satisfies the inequality (16) for all $\sigma \in [0, t]$. From (38), one moreover has the inequality $|f(t - \sigma) - f(t) + \beta\sigma| \leq L\sigma^2/2 + 2N$ for $\sigma = \ell + x$ due to the fact that $f = u - \eta$ and $|\eta(t)| \leq N$. By combining the inequalities and dividing by σ, one obtains

$$\left| \beta - \dot{f}(t) \right| \leq L\sigma + \frac{2N}{\sigma} \tag{39}$$

for $\sigma = \ell + x$. Next, the right-hand side of (39) is shown to be less than or equal to $2\sqrt{2NL} + L\Delta/2$. Since $x \in [0, \Delta]$ and $\Delta \in [0, \ell]$, it follows that

$$0 \geq (x + \Delta/2)(x - \Delta) = x^2 - \frac{\Delta}{2}x - \frac{\Delta^2}{2}$$

$$\geq x^2 - \frac{\Delta}{2}x - \frac{\Delta}{2}\ell = \sigma^2 - 2\ell\sigma + \ell^2 - \frac{\Delta}{2}\sigma. \tag{40}$$

Adding $2\ell\sigma + \Delta\sigma/2$ and multiplying by L/σ then yields

$$2L\ell + L\frac{\Delta}{2} \geq L\sigma + L\frac{\ell^2}{\sigma} \tag{41}$$

which gives $2\sqrt{2NL} + L\Delta/2 \geq L\sigma + 2N/\sigma$ by definition of ℓ. Combining this with (39) establishes the result. ■

Using Lemmata 5.4 and 5.11, an upper bound for the worst-case differentiation error is now proven.

Proposition 5.13. *Let $L, \epsilon > 0$ and consider the differentiator $\mathcal{D}$ defined by (18) with parameters $\overline{\gamma} \geq 1$ and $\overline{T} \in \mathbb{R}_{>0} \cup \{\infty\}$. Let $\overline{N} = L\overline{T}^2/2$ and define the function $\mathcal{T} : \mathbb{R}_{\geq 0} \times [0, \overline{N}) \to \mathbb{R}_{>0}$ as*

$$\mathcal{T}(R, N) = \max\left\{\sqrt{2N/L}, \epsilon\right\}. \tag{42}$$

Then, $M_N^{L,R}[\mathcal{T}(R, N)] \leq \max\{2\sqrt{2}, \overline{\gamma} + \overline{\gamma}^{-1}\}\sqrt{NL}$ holds for all noise bounds $N \in (0, \overline{N})$ and all $R \geq 0$.

Proof. Let $N \in (0, \overline{N})$, $R \in \mathbb{R}_{\geq 0}$ and consider a fixed positive $t \geq \mathcal{T}(R, N) \geq \sqrt{2N/L}$. Relation (18b) along with $\hat{N}(t) \leq N$ and $\gamma(t) \leq \overline{\gamma}$ imply $\hat{T}(t) \leq 2\overline{\gamma}\sqrt{N/L}$. Distinguish hence the cases $\hat{T}(t) \in [\sqrt{2N/L}, 2\overline{\gamma}\sqrt{N/L}]$ and $\hat{T}(t) \in [0, \sqrt{2N/L})$.

In the first case, according to Lemma 5.4,

$$\left|y(t) - \dot{f}(t)\right| \leq \frac{2N}{\hat{T}(t)} + \frac{L\hat{T}(t)}{2}$$

$$\leq \max\left\{\frac{2}{\sqrt{2}} + \frac{\sqrt{2}}{2}, \frac{1}{\overline{\gamma}} + \overline{\gamma}\right\}\sqrt{NL} \tag{43}$$

then holds, wherein $\sqrt{2} + \sqrt{2}^{-1} < 2\sqrt{2}$.

Otherwise, $\hat{T}(t) \geq 2\sqrt{\hat{N}(t)/L}$ follows from (18b), because $\hat{T}(t) < \sqrt{2N/L} \leq t$ and hence having both $\hat{T}(t) = t$ and $\hat{T}(t) = \overline{T} > \sqrt{2N/L}$ is impossible, and $\gamma(t) \geq 1$. Now distinguish the cases $\hat{T}(t) > 0$ and $\hat{T}(t) = 0$. For $\hat{T}(t) > 0$, use Lemma 5.9 to find that (38) holds for $\beta = y(t)$ and $\sigma = \sqrt{2N/L}$, because $\overline{T} \geq \sigma \geq \hat{T}(t)$. For $\hat{T}(t) = 0$, conclude that also $\hat{N}(t) = 0$ and deduce from Lemma 5.8 with $\mu = \min\{t, \overline{T}\}$ that (38) holds for $\beta = y(t)$ and all $\sigma \in [0, \sqrt{2N/L}]$. The bound $|y(t) - \dot{f}(t)| \leq 2\sqrt{2NL}$ then follows from Lemma 5.11 setting $\Delta = 0$. ■

5.5. Proof of the main Theorem 5.1

Proof of Theorem 5.1(a). It is clear that the differentiator is well-defined if $\hat{T}(t) > 0$. If $\hat{T}(t) = 0$, then also $\hat{N}(t) = 0$ and hence Q as defined in (18e) satisfies (29). Existence of the limit in (18a) for all $u \in \mathcal{U}$ then follows from Lemma 5.7.

Proof of Theorem 5.1(b). If $N = 0$, then $\hat{T}(t) = 0$ due to Proposition 5.6. Thus, (18a) with $u = f$ implies $y(t) = \dot{f}(t)$, i.e., $M_0^{L,R}(t) = 0$, for all $t > 0$, proving exactness from the beginning over $\mathcal{F}_L$. For robustness, fix $t > 0$ and set $\epsilon = t$ in Proposition 5.13. For all $N \in (0, Lt^2/8]$ with $N \leq L\overline{T}^2/2$, then

$\mathcal{T}(R, N) = t$ in that proposition, and hence

$$\sup_{f \in \mathcal{F}_L^R} Q_N^f(t) \leq M_N^{L,R}(t) + M_0^{L,R}(t) = M_N^{L,R}(t)$$

$$\leq \max\{2\sqrt{2}, \overline{\gamma} + \overline{\gamma}^{-1}\}\sqrt{NL}, \tag{44}$$

which implies $Q^{L,R}(t) = \lim_{N \to 0} \sup_{f \in \mathcal{F}_L^R} Q_N^f(t) = 0$, proving robustness almost from the beginning over $\mathcal{F}_L$.

Proof of Theorem 5.1(c). Note that $\overline{\gamma} + \overline{\gamma}^{-1} \leq 2\sqrt{2}$ holds for $\overline{\gamma} \in [1, 1 + \sqrt{2}]$. Then, for every $t > \sqrt{2N/L}$ find that the inequality $M_N^{L,R}(t) \leq 2\sqrt{2NL}$ holds for all $N \in (0, L\overline{T}^2/4)$ due to Proposition 5.13 with $\epsilon = t$ as well as for $N = 0$ due to Theorem 5.1(b). Consequently, $C_L \leq 2\sqrt{2}$. Equality follows from the fact that $C_L \geq 2\sqrt{2}$ for all causal differentiators according to Proposition 3.10. ■

Proof of Theorem 5.1(d). Theorem 5.1(c) with $\overline{T} = \infty$ yields $M_N^{L,R}(t) \leq 2\sqrt{2NL}$ for all $N \in [0, \infty)$; hence, $\overline{C}_L \leq 2\sqrt{2}$, with equality again due to Proposition 3.10. ■

6. Sample-based differentiation

When only sampled information is available, the worst-case error between any estimate and the true derivative can never be better than if measurements are available continuously over time. Therefore, some degree of error additional to that obtained in the continuous-time case is to be expected. Taking this fact into account, it will be shown in this section that sampled versions of the proposed differentiators achieve best possible accuracy. Differentiators operating on signal samples will henceforth be called *sample-based differentiators*.

In the following, suppose that only sampled measurements of $u = f + \eta$ are available, at times $t_k = k\Delta$, where $\Delta > 0$ is the sampling period and $k \in \mathbb{N}_0$. Such a differentiator, which takes the samples $u(t_j)$ from $j = 0$ to $j = k$ as input to produce the estimate y_k of the derivative $\dot{f}(t_k)$ is hereafter denoted by $\mathcal{D}_\Delta$, so that[2] $[\mathcal{D}_\Delta u](t_k) = y_k$.

6.1. Worst-case error

The worst-case error in Definition 2.2 can be straightforwardly adapted to the sampled case, as follows.

Definition 6.1. Let $L, N \in \mathbb{R}_{\geq 0}$ and $\Delta > 0$. A sample-based differentiator $\mathcal{D}_\Delta$ is said to have worst-case error $M_N^{L,R}(t)$ from time $t = k\Delta$, $k \in \mathbb{N}_0$, over the signal class $\mathcal{F}_L^R$ with noise bound N if

$$M_N^{L,R}(k\Delta) = \sup_{\substack{u = f + \eta \\ \eta \in \mathcal{E}_N \\ f \in \mathcal{F}_L^R}} \sup_{\substack{\ell \in \mathbb{N}_0 \\ \ell \geq k}} \left|\dot{f}(\ell\Delta) - [\mathcal{D}_\Delta u](\ell\Delta)\right|. \tag{45}$$

It is worth noting that for a sample-based differentiator $\mathcal{D}_\Delta$ the worst-case error $M_N^{L,R}$ is also only defined at integer multiples of the sampling time Δ. As in the continuous-time case, $M_N^{L,R}(k\Delta)$ is non-increasing with respect to $k \in \mathbb{N}_0$ and non-decreasing with respect to $N, L, R \in \mathbb{R}_{\geq 0}$.

6.2. Performance limits and quasi-exactness

A lower bound on the additional error introduced by sampling is given in the following result.

---

[2] Formally, a sample-based differentiator, $\mathcal{D}_\Delta$ with sampling time $\Delta > 0$ is an operator $\mathcal{D}_\Delta : \mathcal{U} \to (\Delta \cdot \mathbb{N}_0 \to \mathbb{R})$, with the additional property that $\mathcal{D}_\Delta(u_1 + u_2) = \mathcal{D}_\Delta(u_1)$ whenever $u_2(k\Delta) = 0$ for all $k \in \mathbb{N}_0$.
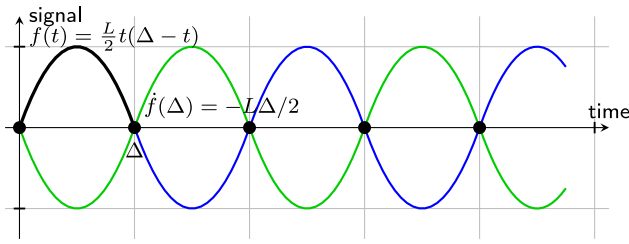
Fig. 3. Black: parabola arc satisfying $\ddot{f}(t) \equiv -L$, $f(0) = 0$ and $f(\Delta) = 0$. Blue and green: signals constructed by piecing together shifted (and sign-changed) copies of the black parabola arc. Black circles: measurements (all zero). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Proposition 6.2. *Let $L, N, R \in \mathbb{R}_{\geq 0}$, $\Delta > 0$. Then, the worst-case error of any sample-based differentiator $\mathcal{D}_\Delta$ satisfies*

$$M_N^{L,R}(k\Delta) \geq \frac{L\Delta}{2}, \quad \forall k \in \mathbb{N}. \tag{46}$$

Remark 6.3. An illustration of the proof is shown in Fig. 3. Essentially, piecing together parabola arcs with second derivative alternately equal to $\pm L$ on time intervals of length Δ yields zero measurements, making two functions with maximum derivative $\pm L\Delta/2$ indistinguishable to the differentiator from the sampled measurements alone. ∘

Proof. The fact that $M_N^{L,R}(k\Delta) \geq M_0^{L,0}(k\Delta)$ follows directly from the definition, hence it suffices to show $M_0^{L,0}(k\Delta) \geq \frac{L\Delta}{2}$. Assume to the contrary that there exist $k_0 \in \mathbb{N}$ and $\epsilon > 0$ such that $M_0^{L,0}(k\Delta) < (1-\epsilon)\frac{L\Delta}{2}$ for $k = k_0$. By Definition 6.1, this is then true also for all integers $k \geq k_0$.

Consider the function $g_{a,b,c} : [0, \Delta] \to \mathbb{R}$ with parameters $a, b \in [0, 1]$ and $c \in [0, \frac{1}{4}]$, defined as

$$g_{a,b,c}(t) = \begin{cases} a\frac{L\Delta}{2}t + \frac{Lt^2}{2} & t \in [0, c\Delta) \\ b\frac{L\Delta^2}{8} - \frac{L}{2}(t - \frac{\Delta}{2})^2 & t \in [c\Delta, \frac{\Delta}{2}) \\ b\frac{L}{2}t(\Delta - t) & t \in [\frac{\Delta}{2}, \Delta]. \end{cases} \tag{47}$$

For arbitrary $a \in [0, 1]$, $b = 1 - \frac{1}{2}(1-a)^2$, and $c = \frac{1}{4}(1-a)$, one verifies that $|\ddot{g}_{a,b,c}(t)| \leq L$ almost everywhere on $[0, \Delta]$, that $g_{a,b,c}(0) = g_{a,b,c}(\Delta) = 0$, that $\dot{g}_{a,b,c}(0) = a\frac{L\Delta}{2}$, and that $\dot{g}_{a,b,c}(\Delta) = -b\frac{L\Delta}{2}$. In particular, $g_{1,1,0}(t) = \frac{L}{2}t(\Delta - t)$ is the black parabola arc depicted in Fig. 3.

Recursively define the (strictly increasing) sequence (a_j) via $a_{j+1} = 1 - \frac{1}{2}(1-a_j)^2$ with $a_0 = 0$, and set $c_j = (1-a_j)/4$, $b_j = a_{j+1}$ for all $j \in \mathbb{N}_0$. Using these sequences, define the function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ piece-wise as

$$f(t) = (-1)^j g_{a_j, b_j, c_j}(t - j\Delta) \quad \text{for } t \in [j\Delta, (j+1)\Delta). \tag{48}$$

From the properties of $g_{a,b,c}$ above, it follows that $f \in \mathcal{F}_L^0$ with $f(k\Delta) = 0$ for all $k \in \mathbb{N}_0$. Applying the inputs $u = f$ and $u = -f$ to the differentiator then produces identical (zero) measurements at the sampling time instants. Hence, $M_0^{L,0}(k\Delta) \geq |\dot{f}(k\Delta)| = a_k\frac{L\Delta}{2}$. Since $\lim_{j \to \infty} a_j = 1$, it is possible to select $k \geq k_0$ such that $a_k \geq 1 - \epsilon$, yielding a contradiction to $M_0^{L,0}(k\Delta) < (1-\epsilon)\frac{L\Delta}{2}$. ∎

As the previous result shows, it is clearly impossible to obtain an exact differentiator based on sampled measurements. Arguably, the closest property to exactness in the sampled case is to achieve equality in (46). This property is called quasi-exactness in the following. Its formal definition is similar to that of exactness in Definition 2.3.

Definition 6.4. A sample-based differentiator $\mathcal{D}_\Delta$ is said, over the signal class $\mathcal{F}_L$, to be

- quasi-exact in finite time if for every R there exists $k_R \in \mathbb{N}$ such that $M_0^{L,R}(k_R\Delta) = \frac{L\Delta}{2}$;
- quasi-exact in fixed time if there exists $k \in \mathbb{N}$ such that $M_0^{L,R}(k\Delta) = \frac{L\Delta}{2}$ for all $R \geq 0$;
- quasi-exact from the beginning if $M_0^{L,R}(\Delta) = \frac{L\Delta}{2}$ for all $R \geq 0$;
- not quasi-exact if it is not quasi-exact in finite time.

A quasi-exact sample-based differentiator yields the best possible estimate for the derivative in the noise-free case based on the available samples. It is worth noting, however, that the discretization of an exact continuous-time differentiator does not necessarily yield a quasi-exact sample-based differentiator.

The simplest quasi-exact differentiator is the first-order difference $y_k = [u(t_k) - u(t_k - \Delta)]/\Delta$, as can be seen from Lemma 4.3 setting $N = 0$. In the presence of noise, however, the worst-case error of this simple differentiator may become prohibitively large. It is hence desirable to achieve a worst-case error that is close to the lower bound of all causal quasi-exact differentiators, which is stated in the following proposition that is analogous to Proposition 3.3.

Proposition 6.5. *Let $\Delta > 0$, $L \geq 0$ and consider a causal sample-based differentiator $\mathcal{D}_\Delta$. Suppose that $\mathcal{D}_\Delta$ is quasi-exact in finite-time over $\mathcal{F}_L$. Then,*

$$M_N^{L,R}(r\Delta) \geq 2\sqrt{2NL} - \frac{L\Delta}{2} \tag{49}$$

holds for all $R, N \geq 0$ and $r \in \mathbb{N}$.

Remark 6.6. Sample-based differentiators are trivially limited also by the bound $M_N^{L,R}(k\Delta) \geq 2\sqrt{NL}$ from Proposition 3.1, which applies to every causal differentiator. Hence, the bound (49) is nontrivial only if $\Delta \leq 4(\sqrt{2} - 1)\sqrt{N/L}$. ∘

Proof. For $L = 0$, (49) is trivial. For $L > 0$, let $k \in \mathbb{N}$ be such that $M_0^{L,0}(k\Delta) = L\Delta/2$ according to Definition 6.4. Furthermore, let $N \geq 0$, $r \in \mathbb{N}$, and define $\kappa = \sqrt{N/L}$, select $\ell \in \mathbb{N}$ such that $\ell \geq \max\{r, k\}$ and $\ell\Delta \geq (2 + \sqrt{2})\kappa$. Define also $T = \ell\Delta$ and $\tau = \ell\Delta - (2 + \sqrt{2})\kappa$. For these values of κ, τ, T, consider the functions g_1 and g_2 in (12). Then, $g_1 \in \mathcal{F}_L^0$ and $g_2 \in \mathcal{E}_N$. Choosing $f = -g_1$ and $\eta = g_2$ yields differentiator input $u(t) = g_1(t)$ for all $t \in [0, T]$. Since $g_1 \in \mathcal{F}_L^0$, then

$$\frac{L\Delta}{2} \geq M_0^{L,0}(k\Delta) \geq M_0^{L,0}(T) \geq |[\mathcal{D}_\Delta u](T) - \dot{g}_1(T)|$$

$$\geq 2|\dot{g}_1(T)| - |[\mathcal{D}_\Delta u](T) + \dot{g}_1(T)|. \tag{50}$$

Since $\dot{f}(T) = -\dot{g}_1(T) = -\sqrt{2NL}$, and $M_N^{L,R}(j\Delta)$ is non-decreasing with respect to R and non-increasing with respect to j, this yields

$$M_N^{L,R}(r\Delta) \geq M_N^{L,R}(T) \geq |[\mathcal{D}_\Delta u](T) - \dot{f}(T)|$$

$$= |[\mathcal{D}_\Delta u](T) + \dot{g}_1(T)| \geq 2\sqrt{2NL} - \frac{L\Delta}{2}, \tag{51}$$

as claimed. ∎

7. Sample-based optimal robust exact differentiation

In this section, a sampled version of the proposed optimal robust exact differentiator is shown.

7.1. Proposed sample-based differentiator

The sampled version of (18) is given by the noise amplitude estimation over a time window of length $\overline{T} = \overline{k}\Delta$ with parameter[3] $\overline{k} \in \mathbb{N} \setminus \{1\}$ according to $\hat{N}_0 = \hat{N}_1 = 0$ and

$$\hat{N}_k = \frac{1}{2} \max_{\substack{\ell \in \{2, \ldots, \overline{k}\} \\ \ell \leq k \\ j \in \{1, \ldots, \ell\}}} \left(|Q(t_k, \ell\Delta, j\Delta)| - \frac{L\Delta^2 j(\ell - j)}{2} \right), \quad (52a)$$

for $k \geq 2$, the selection of γ_k according to[4]

$$\gamma_k \in \mathcal{G}_k \cap [1, \overline{\gamma}], \quad \text{with } \overline{\gamma} \geq 2 \text{ and} \quad (52b)$$

$$\mathcal{G}_k = \begin{cases} \{1\} & \text{if } 2\sqrt{\frac{\hat{N}_k}{L}} \leq \Delta, \\ \left\{ \frac{j\Delta}{2\sqrt{\hat{N}_k/L}} : j \in \mathbb{N} \right\} & \text{otherwise,} \end{cases} \quad (52c)$$

and the computation of the differentiator output via

$$\hat{T}_k = \min \left\{ t_k, \overline{k}\Delta, \max \left\{ \Delta, 2\gamma_k\sqrt{\frac{\hat{N}_k}{L}} \right\} \right\} \quad (52d)$$

$$y_k = \frac{u(t_k) - u(t_k - \hat{T}_k)}{\max\{\hat{T}_k, \Delta\}} \quad (52e)$$

with Q in (18e). As will be shown, the restriction $\overline{\gamma} \geq 2$ ensures that the set $\mathcal{G}_k \cap [1, \overline{\gamma}]$ is always non-empty and that, for $k \geq 1$, $\hat{T}_k/\Delta \in \mathbb{N}$ and $\hat{T}_k = \Delta$ if and only if $\mathcal{G}_k = \{1\}$. From Lemma 5.5, it follows that $\hat{N}_k \leq N$ in analogy to Proposition 5.6.

It will be shown that the sample-based implementation (52) of the optimal robust exact differentiator in (18) is quasi-exact, and that, with appropriately chosen $\overline{\gamma}$, its worst-case differentiation error $M_N^{L,R}(k\Delta)$ is always contained in a band $2\sqrt{2NL} \pm L\Delta/2$ around its optimal continuous-time value. Considering Propositions 6.2 and 6.5, one can see that this is the tightest band of this form one can hope to obtain for a quasi-exact sample-based differentiator. The following main theorem for the sampled case formally states these results. The proof is given in Section 7.4.

Theorem 7.1. *Let $L, \Delta \in \mathbb{R}_{>0}$, $N \in \mathbb{R}_{\geq 0}$, and consider the sample-based differentiator $\mathcal{D}_\Delta$ as given in (52) with parameters $\overline{\gamma} \geq 2$ and $\overline{k} \in \mathbb{N} \setminus \{1\}$. Let $\overline{N} = L\Delta^2(\overline{k} - 1)^2/2$. Then, the following statements are true:*

(a) $\mathcal{D}_\Delta$ is a well-defined sample-based differentiator; specifically, $y_k = 0$ for $k = 0$, and, for all $k \in \mathbb{N}$, the set $\mathcal{G}_k \cap [1, \overline{\gamma}]$ is non-empty and $\hat{T}_k/\Delta \in \mathbb{N}$.

(b) $\mathcal{D}_\Delta$ is quasi-exact from the beginning over $\mathcal{F}_L$.

(c) If $\overline{\gamma} \in [2, 1 + \sqrt{2}]$, then

$$2\sqrt{2NL} - \frac{L\Delta}{2} \leq M_N^{L,R}(k\Delta) \leq 2\sqrt{2NL} + \frac{L\Delta}{2} \quad (53)$$

for all $N \leq \overline{N}$ and for all $k \in \mathbb{N}$ with $k\Delta \geq 2\sqrt{N/L}$. In particular, $M_N^{L,R}(k\Delta) = 2\sqrt{2NL} + \frac{L\Delta}{2}$ for $N = 0$ and all $k \in \mathbb{N}$.

(d) If $\overline{\gamma} \in [2, 3/\sqrt{2})$, then $M_N^{L,R}(k\Delta) = 2\sqrt{2NL} - \frac{L\Delta}{2}$ holds for $N = L\Delta^2/2$ and all $k \in \mathbb{N}$ with $k \geq 2$. △

Remark 7.2 (*Tuning*). Analogous to Remark 5.2, the proposed sample-based differentiator may be tuned using a (crude) noise amplitude upper bound $\overline{N}$ by selecting the smallest integer $\overline{k}$ satisfying $\overline{k}\Delta > \sqrt{2\overline{N}/L} + \Delta$, setting $\overline{\gamma} = 2$, and choosing the smallest possible value for γ_k in every sampling step. ○

---

[3] Like in the continuous-time case, all formal results also hold with infinite window length, i.e., with $\overline{k} = \infty$.

[4] In practice, the smallest element of $\mathcal{G}_k \cap [1, \overline{\gamma}]$ may be chosen.

Remark 7.3. The theorem reveals that, for fixed L, the upper differentiation error bound has the same asymptotic behavior of the order $\max\{\sqrt{N}, \Delta\}$ with respect to sampling period Δ and noise amplitude N as existing robust exact differentiators with sampled measurements, cf. Livne and Levant (2014, Theorems 1 and 3). ○

7.2. Worst-case error upper bound

The following result gives an upper bound on the worst-case error similar to the one in Proposition 5.13.

Proposition 7.4. *Let $L, \Delta \in \mathbb{R}_{>0}$, $u = f + \eta$ with $f \in \mathcal{F}_L$, $\eta \in \mathcal{E}_N$, and consider the sample-based differentiator $\mathcal{D}_\Delta$ as given in (52) with parameters $\overline{\gamma} \geq 2$ and $\overline{k} \in \mathbb{N} \setminus \{1\}$. Define $\overline{N} = L\Delta^2(\overline{k} - 1)^2/2$. Then, the bound $M_N^{L,R}(k\Delta) \leq \max\{2\sqrt{2}, \overline{\gamma} + \overline{\gamma}^{-1}\}\sqrt{NL} + L\frac{\Delta}{2}$ holds for all $N \in [0, \overline{N}]$ and all $k \in \mathbb{N}$ with $k\Delta \geq \sqrt{2N/L}$.*

Proof. Let $t_k = k\Delta$, $N \in [0, \overline{N}]$ and distinguish cases $N \leq L\hat{T}_k^2/2$ and $N > L\hat{T}_k^2/2$. In the first case, $\hat{T}_k \geq \sqrt{2N/L}$ holds. With the inequality $\hat{N}_k \leq N$, which follows from the definition of $\hat{N}_k$ and Lemma 5.5, moreover either $\hat{T}_k \leq 2\gamma_k\sqrt{N/L}$ or $\hat{T}_k = \Delta$ holds. According to Lemma 5.4,

$$\left| y_k - \dot{f}(t_k) \right| \leq \frac{2N}{\hat{T}_k} + \frac{L\hat{T}_k}{2} \quad (54)$$

and if $\hat{T}_k \leq 2\gamma_k\sqrt{N/L}$, then

$$\frac{2N}{\hat{T}_k} + \frac{L\hat{T}_k}{2} \leq \max\left\{ \frac{2}{\sqrt{2}} + \frac{\sqrt{2}}{2}, \frac{1}{\gamma_k} + \gamma_k \right\}\sqrt{NL}$$
$$\leq \max\{2\sqrt{2}, \overline{\gamma} + \overline{\gamma}^{-1}\}\sqrt{NL}. \quad (55)$$

holds. If $\hat{T}_k = \Delta$, then

$$\frac{2N}{\hat{T}_k} + \frac{L\hat{T}_k}{2} \leq \frac{2N}{\sqrt{2N/L}} + L\frac{\Delta}{2} = \sqrt{2NL} + L\frac{\Delta}{2}. \quad (56)$$

In the second case, we have $\hat{T}_k < \sqrt{2N/L} \leq t_k$. Then, $\hat{T}_k \geq 2\sqrt{\hat{N}_k/L}$ follows from (52d), because $\hat{T}_k = t_k$ and $\hat{T}_k = \overline{k}\Delta$ are impossible due to $(\overline{k} - 1)\Delta \geq \sqrt{2N/L}$, because $\hat{T}_k = \Delta$ implies $\mathcal{G}_k = \{1\}$ and $\Delta \geq 2\sqrt{\hat{N}_k/L}$, and because $\gamma_k \geq 1$. Define $\ell = \sqrt{2N/L}$ and let $x \in [0, \Delta)$ be such that $\hat{\sigma} := \ell + x$ satisfies $\hat{\sigma}/\Delta \in \mathbb{N}$. Then, moreover $\hat{\sigma} \leq \sqrt{2N/L} + \Delta \leq \overline{k}\Delta$, and by definition of $\hat{N}_k$ in (52a),

$$\left| Q(t_k, \hat{\sigma}, \hat{T}_k) \right| \leq 2\hat{N}_k + \frac{L\hat{T}_k(\hat{\sigma} - \hat{T}_k)}{2} \quad (57)$$

holds because $\hat{T}_k \leq \hat{\sigma} \leq t_k$ and $\hat{\sigma}/\Delta \in \mathbb{N} \cap [1, \overline{k}]$. Using Lemma 5.9, then (33) holds with $t = t_k$ and $\sigma = \hat{\sigma}$. The result then follows from Lemma 5.11. ∎

7.3. Worst-case error lower bound

Proposition 7.4 shows that for $N = 0$, the worst-case error bound is equal to $L\Delta/2$, corresponding to the noise-free case. This noise-free bound is tight and cannot be improved, as is shown in Proposition 6.2.

For all $N \geq 0$, a lower bound on the worst-case error is given by Proposition 6.5. The following auxiliary lemma will be used to show that the proposed sample-based differentiator can actually attain this lower bound in some specific cases.

Lemma 7.5. *Let $L \in \mathbb{R}_{\geq 0}$, $\Delta > 0$ and $N = L\Delta^2/2$. Then, the sample-based differentiator $\mathcal{D}_\Delta$ given in (52) with parameters $\overline{\gamma} \geq 2$ and $\overline{k} \in \mathbb{N} \setminus \{1\}$ achieves*

$$M_N^{L,R}(t_k) \leq 2\sqrt{2NL} - \frac{L\Delta}{2} \tag{58}$$

whenever $t_k \geq \hat{T}_k$ and $\hat{T}_k \in \{\Delta, 2\Delta\}$.

Proof. According to Lemma 5.4 and provided $t_k \geq \hat{T}_k$, it follows that

$$\left|[\mathcal{D}_\Delta u](t_k) - \dot{f}(t_k)\right| \leq \frac{2N}{\hat{T}_k} + \frac{L\hat{T}_k}{2} = \frac{3}{2}L\Delta$$
$$= 2\sqrt{2NL} - \frac{L\Delta}{2} \tag{59}$$

for the considered value $N = L\Delta^2/2$. ∎

7.4. Proof of the main Theorem 7.1

Proof of Theorem 7.1(a). For $k = 0$, $\hat{T}_k = t_k = 0$ and hence $y_k = 0$. For $k \in \mathbb{N}$, if $\hat{N}_k \leq L\Delta^2/4$, then $\mathcal{G}_k \cap [1, \overline{\gamma}] = \{1\}$; hence $\gamma_k = 1$ and $\hat{T}_k = \Delta$. Otherwise, $\Delta/\sqrt{4\hat{N}_k/L} \leq 1$, and thus the difference between consecutive elements in $\mathcal{G}_k$ is at most one, proving non-emptiness of $\mathcal{G}_k \cap [1, \overline{\gamma}]$ with $\overline{\gamma} \geq 2$. In this case, $\gamma_k = j\Delta/\sqrt{4\hat{N}_k/L}$ for some j, and hence $\hat{T}_k/\Delta \in \{k, \overline{k}, j\} \subset \mathbb{N}$. ∎

Proof of Theorem 7.1(b). Let $R \geq 0$ and set $N = 0$ in Proposition 7.4 to obtain $M_0^{L,R}(k\Delta) \leq L\Delta/2$ for all $k \in \mathbb{N}$. Hence, $M_0^{L,R}(\Delta) = L\Delta/2$ by Proposition 6.2. ∎

Proof of Theorem 7.1(c). The upper bound follows from Proposition 7.4 and the fact that $\overline{\gamma} + \overline{\gamma}^{-1} \leq 2\sqrt{2}$ for all $\overline{\gamma} \in [2, 1+\sqrt{2}]$. For $N = 0$, equality to this upper bound follows from Proposition 6.2. The lower bound is a consequence of Proposition 6.5 and quasi-exactness shown in Theorem 7.1(b).

Proof of Theorem 7.1(d). For $\overline{\gamma} \in [2, 3/\sqrt{2})$ and with $N = L\Delta^2/2$, one has

$$\hat{T}_k \leq 2\gamma_k\sqrt{\frac{\hat{N}_k}{L}} \leq 2\overline{\gamma}\sqrt{\frac{N}{L}} = \overline{\gamma}\sqrt{2}\Delta < 3\Delta. \tag{60}$$

Hence $\hat{T}_k \in \{\Delta, 2\Delta\}$ and $k\Delta \geq 2\Delta \geq \hat{T}_k$. Lemma 7.5 then yields equality to the lower bound given in Theorem 7.1(c), i.e., $M_N^{L,R}(k\Delta) = 2\sqrt{2NL} - \frac{L\Delta}{2}$. ∎

8. Simulation results

For illustration purposes, the proposed optimal robust exact differentiator (52) is compared to the sliding-mode based RED (6). A signal $f \in \mathcal{F}_L^R$ to be differentiated is chosen as the parabola arc $f(t) = Lt^2/2 + Rt$. The differentiators are implemented in discrete time with sampling period $\Delta = 0.01$. For the RED, the implicit discretization described in Mojallizadeh et al. (2021a) is used with differentiator output $y_k = y_{2,k+1}$. Two different parameter settings are used for the RED: $\lambda_1 = 1.5$, $\lambda_2 = 1.1$ as suggested by Levant (1998) and $\lambda_1 = 2r$, $\lambda_2 = r^2$ obtained using the tuning procedure used in the toolbox by Andritsch et al. (2021) with a robustness factor of $r = 1.4$. For the proposed differentiator, the discrete-time realization in (52) is used, where the smallest possible value for $\gamma_k \in \mathcal{G}_k \cap [1, 2]$ is selected at every sampling time instant. Signal and noise parameters are chosen as $L = R = 1$ and $N = 0.08$. Computational complexity is limited by selecting $\overline{k} = 200$, corresponding to a continuous-time window length
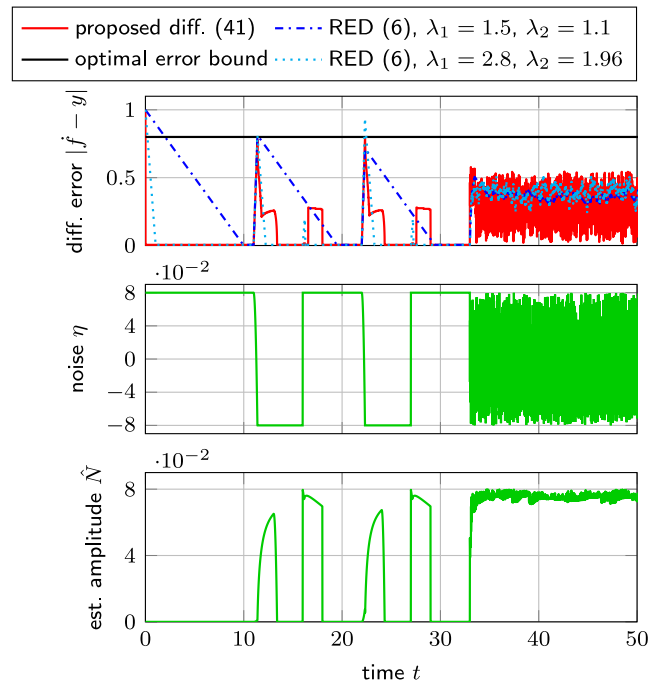
Fig. 4. Simulation results comparing the RED and the proposed optimal robust exact differentiator with $L = 1$, $N = 0.08$, and sampling time $\Delta = 0.01$. The plots show the differentiation error $|\dot{f}(t_k) - y_k|$ (top), noise signal $\eta(t_k)$ (center), and noise amplitude estimate $\hat{N}_k$ (bottom) obtained by the proposed differentiator. The optimal error bound is $2\sqrt{2NL} = 0.8$. Maximum errors for $t \geq 10$ of the proposed differentiator, the RED with $\lambda_1 = 1.5$, and the RED with $\lambda_1 = 2.8$ are 0.7939, 0.8135, and 0.9374, respectively.

$\overline{T} = \overline{k}\Delta = 2$ and guaranteeing optimal performance for noise amplitudes up to $\overline{N} = L\Delta^2(\overline{k}-1)^2/2 \approx 1.98$.

Fig. 4 shows the simulation results. The noise signal is shown in the center portion. Motivated by Fig. 2 and practical considerations, it consists of constant segments, two parabola arcs of the principal form $N - (1 + \lambda_2)Lt^2/2$, with λ_2 taken from the two RED parameter sets, two step jumps from $-N$ to N, and white noise obtained by sampling a uniformly distributed random number from the interval $[-N, N]$. The top and bottom portions of Fig. 4 depict the differentiation error $|\dot{f}(t_k) - y_k|$ and the noise amplitude estimate $\hat{N}_k$.

Initially, the noise is constant and the differentiators hence behave as when differentiating a noise-free signal. One can see that all differentiators make an error equal to R initially; the proposed differentiator then immediately attains quasi-exactness with error bounded by $L\Delta/2$ after a single sampling step, while the RED exhibits a finite convergence time, which decreases with increasing λ_2.

The parabola arcs in the noise lead to peaks in the differentiation error. With the proposed differentiator, they stay below the optimal worst-case error $2\sqrt{2NL}$. Each RED, in presence of the noise parabola constructed using its respective value of λ_2, makes an error that exceeds the optimal error bound, with larger values of λ_2 leading to larger errors. Tuning of the RED thus requires a tradeoff between worst-case convergence speed and worst-case error, while the proposed optimal robust exact differentiator achieves instant convergence with least possible worst-case error bound.

In the presence of white noise, finally, the frequent step-wise changes in the noise allow for a very accurate estimation of the noise amplitude as predicted by Proposition 5.6. However, the proposed differentiator also exhibits the largest variation in the error, due to its direct feed-through of the noisy input and the

resulting absence of any noise filtering. The RED, on the other hand, filters high-frequency components of the noise, leading to a smoother error signal overall. Nevertheless, all differentiators lead to an error of similar magnitude in this case.

9. Conclusion

A first-order differentiator that is robust and exact over a wide class of signals and that achieves optimal differentiation accuracy is proposed for the first time. It is based on the structure of a linear differentiator with a parameter that adapts based on a suitable estimate of the noise amplitude. It is shown that, in the presence of noise, the proposed differentiator achieves the lowest possible worst-case error among all exact differentiators, and that it converges instantaneously to the true derivative in the absence of noise, hence outperforming all fixed-time convergent differentiators in terms of convergence speed.

For the sampled-data case, a discrete-time implementation of the differentiator based on sampled measurements is provided. This sample-based differentiator is shown to retain the properties of the optimal continuous-time differentiator in their closest possible forms. In the absence of noise, the proposed sample-based differentiator has the property that it has the least possible worst-case error among all sample-based differentiators; this property is hence called quasi-exactness. It moreover attains this quasi-exactness after the least possible convergence time of a single sampling step. In the presence of noise, the worst-case error is shown to converge to a band around its continuous-time optimal value, whose width is as small as possible and is a linear function of the sampling time.

Acknowledgement

References

Andritsch, B., Horn, M., Koch, S., Niederwieser, H., Wetzlinger, M., & Reichhartinger, M. (2021). The robust exact differentiator toolbox revisited: Filtering and discretization features. In *IEEE international conference on mechatronics (ICM)*.

Barbot, J.-P., Levant, A., Livne, M., & Lunz, D. (2020). Discrete differentiators based on sliding modes. *Automatica, 112*, 108633.

Bhat, S. P., & Bernstein, D. S. (2000). Finite-time stability of continuous autonomous systems. *SIAM Journal on Control and Optimization, 38*(3), 751–766.

Carvajal-Rubio, J. E., Sánchez-Torres, J. D., Defoort, M., Djemai, M., & Loukianov, A. G. (2021). Implicit and explicit discrete-time realizations of homogeneous differentiators. *International Journal of Robust and Nonlinear Control, 31*, 3606–3630.

Cruz-Zavala, E., Moreno, J. A., & Fridman, L. (2011). Uniform robust exact differentiator. *IEEE Transactions on Automatic Control, 56*(11), 2727–2733.

Filippov, A. F. (1988). *Differential equations with discontinuous right-hand side*. Dordrecht, The Netherlands: Kluwer Academic Publishing.

Fraguela, L., Angulo, M. T., Moreno, J. A., & Fridman, L. (2012). Design of a prescribed convergence time uniform robust exact observer in the presence of measurement noise. In *Proc. 51st IEEE conf. on decision and control* (pp. 6615–6620). Maui, Hawaii, USA.

Ghanes, M., Barbot, J.-P., Fridman, L., Levant, A., & Boisliveau, R. (2020). A new varying-gain-exponent-based differentiator/observer: An efficient balance between linear and sliding-mode algorithms. *IEEE Transactions on Automatic Control, 65*(12), 5407–5414.

Haimovich, H., Seeber, R., Aldana-López, R., & Gómez-Gutiérrez, D. (2022). Differentiator for noisy sampled signals with best worst-case accuracy. *IEEE Control Systems Letters, 6*, 938–943.

Hanan, A., Levant, A., & Jbara, A. (2021). Low-chattering discretization of homogeneous differentiators. *IEEE Transactions on Automatic Control, 67*, 2946–2956.

Khalil, H. K., & Praly, L. (2014). High-gain observers in nonlinear feedback control. *International Journal of Robust and Nonlinear Control, 24*, 993–1015.

Kolmogorov, A. N. (1962). On inequalities between the upper bounds of the successive derivatives of an arbitrary function on an infinite interval. *American Mathematical Society Translations Series 1, 2*, 233–243.

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

Levant, A. (2003). Higher-order sliding modes, differentation and output-feedback control. *International Journal of Control, 76*(9/10), 924–941.

Levant, A. (2007). Finite differences in homogeneous discontinuous control. *IEEE Transactions on Automatic Control, 52*(7), 1208–1217.

Levant, A. (2012). Digital sliding-mode-based differentiation. In *12th IEEE workshop on variable structure systems, VSS'12*. Mumbai, India.

Levant, A., & Livne, M. (2020). Robust exact filtering differentiators. *European Journal of Control, 55*, 33–44.

Levant, A., Livne, M., & Yu, X. (2017). Sliding-mode-based differentiation and its application. In *20th IFAC world congress* (pp. 1699–1704). Elsevier.

Levant, A., & Yu, X. (2018). Sliding-mode-based differentiation and filtering. *IEEE Transactions on Automatic Control, 63*(9), 3061–3067.

Livne, M., & Levant, A. (2014). Proper discretization of homogeneous differentiators. *Automatica, 50*, 2007–2014.

Mboup, M., Join, C., & Fliess, M. (2009). Numerical differentiation with annihilators in noisy environment. *Numerical Algorithms, 50*, 439–467.

Mojallizadeh, M. R., Brogliato, B., & Acary, V. (2021a). Time-discretizations of differentiators: Design of implicit algorithms and comparative analysis. *International Journal of Robust and Nonlinear Control, 31*(16), 7679–7723.

Mojallizadeh, M. R., Brogliato, B., & Acary, V. (2021b). Time-discretizations of differentiators: Design of implicit algorithms and comparative analysis. *International Journal of Robust and Nonlinear Control, 31*, 7679–7723.

Schoenberg, I. J., & Cavaretta, A. (1970). *Solution of Landau's problem concerning higher derivatives on the halfline*: *Technical summary report*, Madison, Wisconsin: University of Wisconsin–Madison, Mathematics Research Center.

Seeber, R., Haimovich, H., Horn, M., Fridman, L., & De Battista, H. (2021). Robust exact differentiators with predefined convergence time. *Automatica, 134*(109858).

Seeber, R., Horn, M., & Fridman, L. (2018). A novel method to estimate the reaching time of the super-twisting algorithm. *IEEE Transactions on Automatic Control, 63*(12), 4301–4308.

Vasiljevic, L. K., & Khalil, H. K. (2008). Error bounds in differentiation of noisy signals by high-gain observers. *Systems & Control Letters, 57*(10), 856–862.

Yu, X., & Xu, J.-X. (1996). Nonlinear derivative estimator. *Electronics Letters, 32*(16).
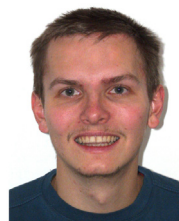
Richard Seeber received his M.Sc. degree in Electrical Engineering from Graz University of Technology in 2012, and he completed his Ph.D. at the Institute of Automation and Control, Graz University of Technology in 2016. He currently holds a Postdoc position at the Christian Doppler Laboratory for Model Based Control of Complex Test Bed Systems. His research interests include theory of sliding mode controllers and differentiators, control of automotive test beds, and control of systems with actuator constraints.

Hernan Haimovich received the Electronics Engineering degree with highest honours in 2001 from the Universidad Nacional de Rosario (UNR), Argentina, and the Ph.D. degree from The University of Newcastle, Australia, in 2006. In 2006, he worked as a Research Assistant at the Centre for Complex Dynamic Systems and Control at the University of Newcastle, Australia, and later as an Argentine Research Council (CONICET) Postdoctoral Research Fellow at the UNR, Argentina. Since 2007, Dr. Haimovich holds a permanent Investigator position from CONICET, currently at the International French-Argentine Center for Information and Systems Science (CIFASIS). Dr. Haimovich is also Professor at the School of Electronics Engineering, UNR. Dr. Haimovich currently serves as Associate Editor for Automatica. His research interests include control theory for nonlinear, switched, time-varying, impulsive, hybrid and/or discontinuous systems, and related applications.