Differentiator for Noisy Sampled Signals with Best Worst-Case Accuracy*

Hernan Haimovich, Richard Seeber, Rodrigo Aldana-López, and David Gómez-Gutiérrez

Abstract—This paper proposes a differentiator for sampled signals with bounded noise and bounded second derivative. It is based on a linear program derived from the available sample information and requires no further tuning beyond the noise and derivative bounds. A tight bound on the worst-case accuracy, i.e., the worst-case differentiation error, is derived, which is the best among all causal differentiators and is moreover shown to be obtained after a fixed number of sampling steps. Comparisons with the accuracy of existing high-gain and sliding-mode differentiators illustrate the obtained results.

Index Terms—Differentiation; Optimization; Estimation; Observers

I. INTRODUCTION

ESTIMATING in real-time the derivatives of a signal affected by noise is a fundamental problem in control theory and continues to be an active area of research, see, e.g., the special issue [1], the comparative analysis [2], and the references therein. Differentiators are often used, for instance, for state estimation [3], Proportional-Derivative controllers, fault detection [4], [5], and unknown input observers [6].

Popular existing methods for differentiation include linear high-gain observers [7], linear algebraic differentiators [8], [9], and sliding mode differentiators [10], [11]. These differ in terms of their convergence properties; while high-gain differentiators converge exponentially [7], algebraic and sliding-mode differentiators exhibit convergence in finite or fixed time [8], [10], and they converge exactly for different classes of noise-free signals [9], [11]. With measurement noise, the accuracy, i.e., the achievable worst-case differentiation error, is limited for all differentiators. Linear differentiators may be tuned to minimize the bound on the differentiation error when the noise amplitude and a bound on the derivative's Lipschitz constant are known [7], whereas the tuning of sliding mode differentiators only requires knowledge about the Lipschitz constant but not about the noise [10], [12].

In practice, differentiation is typically performed on a digital computer using sampled signals. Hence, the use of continuous-time differentiators requires discretization, which is particularly challenging for sliding-mode differentiators because an explicit (forward) Euler discretization may lead to reduced accuracy, numerical chattering, and even instability [13], [14]. Several techniques for that purpose have therefore been proposed, cf. [2], [15], [16]. In any case, the inherent performance limitations of continuous-time differentiators cannot be surpassed in the discrete domain via discretization.

The present paper proposes a differentiator that considers the information available in the sampled signal in the form of a linear program. This approach also yields upper and lower bounds for the derivative, similar to interval observers [17]. Interval observers, however, have seldom been applied to differentiation, see e.g., [18], and they are limited in terms of accuracy by their underlying observer.

In contrast to other observers, the present approach is shown to have the best possible worst-case accuracy among all causal differentiators. This best possible worst-case accuracy is shown to be achieved using a fixed number of samples, thus providing a limit on the computational complexity of the linear program and guaranteeing convergence in a fixed time, similarly to algebraic and some sliding-mode differentiators. Moreover, implementing the algorithm only requires knowledge of the derivative's Lipschitz constant and the noise bound but, unlike other differentiators, yields such an estimate without requiring any further tuning.

Notation: $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$ denote the positive and nonnegative integers, and the reals and nonnegative reals, respectively. If $\alpha \in \mathbb{R}$, then $|\alpha|$ denotes its absolute value. For $x, y \in \mathbb{R}^n$, inequalities and absolute value operate componentwise, so that $|x| \in \mathbb{R}_{\geq 0}^n$ denotes the vector with components $|x_i|$, and $x \leq y$ the set of inequalities $x_i \leq y_i$, for $i = 1, \ldots, n$. For a (differentiable) function $f : D \subset \mathbb{R} \to \mathbb{R}$, $f^{(i)}$ denotes its i-th order derivative. For $a \in \mathbb{R}$, the greatest integer not greater than a is denoted by $\lfloor a \rfloor$. The symbols $\mathbf{0}$, $\mathbf{I}$ and $\mathbf{1}$ denote the zero vector or matrix, the identity matrix and a vector all of whose components are equal to one, respectively.

Hernan Haimovich is with Centro Internacional Franco-Argentino de Ciencias de la Información y de Sistemas (CIFASIS) CONICET-UNR, 2000 Rosario, Argentina. (e-mail: haimovich@cifasis-conicet.gov.ar)

Richard Seeber is with the Christian Doppler Laboratory for Model Based Control of Complex Test Bed Systems, Institute of Automation and Control, Graz University of Technology, Graz, Austria. (e-mail: richard.seeber@tugraz.at)

Rodrigo Aldana-López is with the Department of Computer Science and Systems Engineering, University of Zaragoza, Zaragoza, Spain. (e-mail: rodrigo.aldana.lopez@gmail.com)

David Gómez-Gutiérrez is with Intel Labs, Intel Corporation and with Tecnologico de Monterrey, Jalisco, Mexico. (e-mail: david.gomez.g@ieee.org)

II. PROBLEM STATEMENT AND FORMULATION

A. Problem statement

Consider a differentiable function $\mathfrak{f} : \mathbb{R}_{\geq 0} \to \mathbb{R}$ of which we know that its derivative $\mathfrak{f}^{(1)}$ is globally Lipschitz continuous with Lipschitz constant L, i.e.

$$|\mathfrak{f}^{(2)}(t)| \leq L, \quad \text{for almost all } t \in \mathbb{R}_{\geq 0}. \tag{1}$$

Suppose that a noisy measurement m_k of the value $\mathfrak{f}(kT)$ becomes available at each time instant $t_k = kT$, with $k \in \mathbb{N}_0$, and that a bound N on the noise is known, so that

$$m_j = \mathfrak{f}_j + \eta_j, \qquad \mathfrak{f}_j := \mathfrak{f}(jT), \qquad |\eta_j| \leq N, \tag{2}$$

for $j = 0, 1, \ldots, k$.

The problem to be addressed is to design an algorithm that, at every time instant t_k when a new measurement m_k becomes known, gives all available information on the current value

$$\mathfrak{f}_k^1 := \mathfrak{f}^{(1)}(t_k), \quad t_k := kT, \tag{3}$$

of the derivative of $\mathfrak{f}$. The more specific problem to be solved is as follows.

Problem 1: Devise an algorithm that, given the constants $L \geq 0$ (bound on 2nd-order derivative), $N \geq 0$ (noise bound), $T > 0$ (sampling period) and $k \in \mathbb{N}$, provides all possible values of $\mathfrak{f}_k^1$ based on knowledge of the bound (1) and the measurements m_j for $j = 0, 1, \ldots, k$.

B. Possible values for the derivative

Let $\mathcal{F}_k^1(\mathbf{m}_k)$ denote the set of possible values for $\mathfrak{f}_k^1 = \mathfrak{f}^{(1)}(t_k)$ that are consistent with the bound (1) and the measurements

$$\mathbf{m}_k := [m_0, \ m_1, \ \ldots, \ m_k]^T \tag{4}$$

that satisfy (2). The set $\mathcal{F}_k^1(\mathbf{m}_k)$ can be defined as

$$\mathcal{F}_k^1(\mathbf{m}_k) := \{\mathfrak{f}_k^1 \in \mathbb{R} : \exists \mathfrak{f}(\cdot) \text{ satisfying } (1),(2),(3)\}.$$

The set $\mathcal{F}_k^1(\mathbf{m}_k)$ is convex and hence, whenever nonempty, it will have the form of an interval. Problem 1 can thus be posed as finding the extreme values

$$\overline{\mathcal{F}}_k^1(\mathbf{m}_k) := \sup \mathcal{F}_k^1(\mathbf{m}_k), \quad \underline{\mathcal{F}}_k^1(\mathbf{m}_k) := \inf \mathcal{F}_k^1(\mathbf{m}_k). \tag{5}$$

For future reference, define $\mathcal{X}_k(\mathfrak{f})$ as the vector

$$\mathcal{X}_k(\mathfrak{f}) := [\mathfrak{f}(t_0), \ldots, \mathfrak{f}(t_k), \mathfrak{f}^{(1)}(t_0), \ldots, \mathfrak{f}^{(1)}(t_k)]^T. \tag{6}$$

C. Samples and measurements

Since the derivative $\mathfrak{f}^{(1)}$ is globally Lipschitz continuous, then $\mathfrak{f}^{(2)}$ exists almost everywhere and

$$\mathfrak{f}^{(1)}(\tau) = \mathfrak{f}_j^1 - \int_{\tau}^{t_j} \mathfrak{f}^{(2)}(s)ds, \quad \mathfrak{f}(t) = \mathfrak{f}_j - \int_{t}^{t_j} \mathfrak{f}^{(1)}(\tau)d\tau.$$

From these expressions one can obtain the bounds

$$|\mathfrak{f}_{j-1} - \mathfrak{f}_j + \mathfrak{f}_j^1 T| \leq L\frac{T^2}{2}, \tag{7}$$

$$|\mathfrak{f}_j^1 - \mathfrak{f}_{j-1}^1| \leq LT, \tag{8}$$

and from (2), also $\quad |\mathfrak{f}_j - m_j| \leq N. \tag{9}$

At time $t_k = kT$, every function $\mathfrak{f}$ that satisfies the bound (1) for almost all $t \in [0, t_k]$ will have associated values $\mathcal{X}_k(\mathfrak{f})$ that must satisfy the constraints (7)–(8) for $j = 1, 2, \ldots, k$. In addition, given the noise bound N, the corresponding measurements must satisfy (9) for $j = 0, 1, \ldots, k$.

III. MAIN RESULTS

A. Derivation of the proposed differentiator

Consider a vector $\mathbf{x}_k$ of $2k + 2$ *optimization variables*

$$\mathbf{x}_k := [(f_{0:k})^T, \ (f_{0:k}^1)^T]^T \in \mathbb{R}^{2k+2} \tag{10}$$

$$f_{0:k} := [f_0, \ f_1, \ \ldots, \ f_k]^T, \quad f_{0:k}^1 := [f_0^1, \ f_1^1, \ \ldots, \ f_k^1]^T,$$

where f_i and f_i^1 model possible (hypothetical) values for $\mathfrak{f}_i$ and $\mathfrak{f}_i^1$, respectively. For every $\mathbf{m} \in \mathbb{R}^{k+1}$, consider the set

$$\mathcal{C}_k(\mathbf{m}) := \{\mathbf{x} \in \mathbb{R}^{2k+2} : |A_k\mathbf{x} + M_k\mathbf{m}| \leq \mathbf{b}_k\}, \tag{11}$$

$$A_k = \begin{bmatrix} \mathbf{0} & D_k \\ D_k & -T[\mathbf{0} \ \mathbf{I}] \\ \mathbf{I} & \mathbf{0} \end{bmatrix}, \quad M_k = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{I} \end{bmatrix}, \quad \mathbf{b}_k = \begin{bmatrix} LT\mathbf{1} \\ L\frac{T^2}{2}\mathbf{1} \\ N\mathbf{1} \end{bmatrix}$$

where $D_k \in \mathbb{R}^{k \times k+1}$ is a Toeplitz matrix with first row and column given by $[-1, 1, 0, \ldots, 0]$ and $[-1, 0, \ldots, 0]^T$, respectively. The set $\mathcal{C}_k(\mathbf{m})$ is defined so that whenever a function $\mathfrak{f}$ satisfies (1) and produces the measurements $\mathbf{m}_k$ satisfying (2), then $\mathcal{X}_k(\mathfrak{f}) \in \mathcal{C}_k(\mathbf{m}_k)$. This is so because the rows of the matrices A_k, M_k and the vector $\mathbf{b}_k$ are grouped into 3 blocks, of k, k and $k+1$ rows, where the first block corresponds to (8), the second to (7), and the third to (9).

Remark 1: Although $\mathcal{X}_k(\mathfrak{f}) \in \mathcal{C}_k(\mathbf{m}_k)$ holds for all admissible functions $\mathfrak{f}$ and corresponding measurements $\mathbf{m}_k$, given an arbitrary vector $\mathbf{m}_k \in \mathbb{R}^{k+1}$ with nonempty $\mathcal{C}_k(\mathbf{m}_k)$ it may not be true that a function $\mathfrak{f}$ exists satisfying (1)–(2) with $\mathcal{X}_k(\mathfrak{f}) \in \mathcal{C}_k(\mathbf{m}_k)$ (see the Appendix for a counterexample). The proposed differentiator provides an estimate $\hat{\mathfrak{f}}_k^1$ for the derivative $\mathfrak{f}_k^1 = \mathfrak{f}^{(1)}(t_k)$ by solving the optimization problems (12a)–(12b) and computing (12c):

$$\overline{f}_k^1(\mathbf{m}_k) := \max\{c_k^T\mathbf{x}_k \ : \mathbf{x}_k \in \mathcal{C}_k(\mathbf{m}_k)\}, \tag{12a}$$

$$\underline{f}_k^1(\mathbf{m}_k) := \min\{c_k^T\mathbf{x}_k \ : \mathbf{x}_k \in \mathcal{C}_k(\mathbf{m}_k)\}, \tag{12b}$$

$$\hat{\mathfrak{f}}_k^1 := \left(\overline{f}_k^1(\mathbf{m}_k) + \underline{f}_k^1(\mathbf{m}_k)\right)/2, \tag{12c}$$

with $c_k = [0, \ \ldots, \ 0, \ 1]^T \in \mathbb{R}^{2k+2}$. Note that $c_k^T\mathbf{x}_k = f_k^1$, according to (10). From Remark 1, it follows that the set of possible values for the derivative of $\mathfrak{f}$ at time t_k, namely $\mathcal{F}_k^1(\mathbf{m}_k)$, satisfies $\mathcal{F}_k^1(\mathbf{m}_k) \subseteq [\underline{f}_k^1(\mathbf{m}_k), \overline{f}_k^1(\mathbf{m}_k)]$ and thus

$$\underline{f}_k^1(\mathbf{m}_k) \leq \underline{\mathcal{F}}_k^1(\mathbf{m}_k) \leq \overline{\mathcal{F}}_k^1(\mathbf{m}_k) \leq \overline{f}_k^1(\mathbf{m}_k). \tag{13}$$

The set $\mathcal{C}_k(\mathbf{m}_k)$ is defined by linear inequalities in the optimization variables, for every $\mathbf{m}_k$. Thus, (12a) and (12b) are linear programs; the only information required to implement them are the values L, N, T and the measurements $\mathbf{m}_k$, obtained up to t_k. The proposed estimate $\hat{\mathfrak{f}}_k^1$ yields the smallest worst-case distance to any value within $[\underline{f}_k^1(\mathbf{m}_k), \overline{f}_k^1(\mathbf{m}_k)]$.

The computational complexity of the linear programs increases with increasing k. A fixed number of samples $\hat{K} + 1$ can be considered to limit the complexity as summarized in

Algorithm 1, which is meant to be executed at every time instant. The next section provides a way to choose $\hat{K}$ by studying the worst-case accuracy of the differentiator and showing that a finite K can be computed such that for all $\hat{K} \geq K$ the same worst-case accuracy is obtained.

Algorithm 1: Estimation of $\mathfrak{f}^{(1)}(kT)$, based on $\hat{K}+1$ noisy measurements, using linear programming.

---

input: L, N, T, $\hat{K}$, $\mathbf{m}_k$
Set $\underline{k} := \min\{k, \hat{K}\}$
Set $\mathbf{m}_{\underline{k}}^k := [m_{k-\underline{k}}, \ldots, m_{k-1}, m_k]^T \in \mathbb{R}^{\underline{k}+1}$
Set $A_{\underline{k}}$, $M_{\underline{k}}$, and $\mathbf{b}_{\underline{k}}$ as in (11) using L, N, T.
Set $c_{\underline{k}} := [0, \ldots, 0, 1]^T \in \mathbb{R}^{2\underline{k}+2}$
Solve $\overline{f}_k^1 := \max\left\{c_{\underline{k}}^T \mathbf{x} : |A_{\underline{k}}\mathbf{x} + M_{\underline{k}}\mathbf{m}_{\underline{k}}^k| \leq \mathbf{b}_{\underline{k}}\right\}$
Solve $\underline{f}_{-k}^1 := \min\left\{c_{\underline{k}}^T \mathbf{x} : |A_{\underline{k}}\mathbf{x} + M_{\underline{k}}\mathbf{m}_{\underline{k}}^k| \leq \mathbf{b}_{\underline{k}}\right\}$
return: $\hat{\mathfrak{f}}_k^1 := \left(\overline{f}_k^1 + \underline{f}_{-k}^1\right)/2$

B. Differentiator convergence and worst-case accuracy

A measure of the accuracy of the differentiator is given by the difference between the upper and lower bounds that can be ensured on the derivative

$$w_k(\mathbf{m}_k) := \overline{f}_k^1(\mathbf{m}_k) - \underline{f}_{-k}^1(\mathbf{m}_k). \tag{14}$$

With the differentiator output $\hat{\mathfrak{f}}_k^1$ suggested above, the differentiator error is then bounded by $w_k(\mathbf{m}_k)/2 \geq 0$ if $\mathcal{C}_k(\mathbf{m}_k) \neq \emptyset$. A related quantity is the difference of actual worst-case derivative bounds

$$\mathcal{W}_k(\mathbf{m}_k) := \overline{\mathcal{F}}_k^1(\mathbf{m}_k) - \underline{\mathcal{F}}_k^1(\mathbf{m}_k), \tag{15}$$

which according to (5) correspond to the best possible accuracy obtainable from the measurements $\mathbf{m}_k$.

Let $\mathcal{M}_k(L, N, T)$ denote the set of all possible measurements $\mathbf{m}_k$ that could be obtained for functions satisfying (1) with additive measurement noise bound N:

$$\mathcal{M}_k(L, N, T) := \left\{\mathbf{m}_k \in \mathbb{R}^{k+1} : \exists \mathfrak{f}(\cdot), (1) \text{ and } (2) \text{ hold}\right\}.$$

Consider the obtained and the best possible accuracy over all possible measurements, i.e., their worst-case values,

$$\overline{w}_k(L, N, T) := \sup_{\mathbf{m} \in \mathcal{M}_k(L, N, T)} w_k(\mathbf{m}), \tag{16}$$

$$\overline{\mathcal{W}}_k(L, N, T) := \sup_{\mathbf{m} \in \mathcal{M}_k(L, N, T)} \mathcal{W}_k(\mathbf{m}). \tag{17}$$

Clearly, $\overline{w}_k(L, N, T) \geq \overline{\mathcal{W}}_k(L, N, T)$ according to (13). Also, no causal differentiator can achieve a better worst-case accuracy than $\overline{\mathcal{W}}_k(L, N, T)$ due to (5).

Our main result is the following.

Theorem 1: Given positive L, N, and T, the accuracies $w_k(\mathbf{m})$, $\overline{w}_k(L, N, T)$ obtained with the differentiator (12) and the best possible accuracies $\mathcal{W}_k(\mathbf{m})$, $\overline{\mathcal{W}}_k(L, N, T)$, as defined in (14), (16) and (15), (17), respectively, satisfy:
a) $w_k(\mathbf{0}) = 2\overline{h}_k$, with

$$\overline{h}_k = h_o(\underline{k}), \quad \underline{k} = \min\{k, K\}, \quad h_o(\ell) := \frac{1}{2}LT\ell + \frac{2N}{T\ell},$$

$$K := \begin{cases} Q & \text{if } Q^2 + Q \geq \frac{4N}{LT^2}, \\ Q+1 & \text{otherwise}, \end{cases} \quad Q := \left\lfloor \frac{2}{T}\sqrt{\frac{N}{L}} \right\rfloor.$$

b) $w_k(\mathbf{0}) \geq w_{k+1}(\mathbf{0})$ for all $k \in \mathbb{N}$;
c) $w_k(\mathbf{0}) = \mathcal{W}_k(\mathbf{0})$;
d) $w_k(\mathbf{0}) = \overline{w}_k(L, N, T) = \overline{\mathcal{W}}_k(L, N, T)$.

Items a) and b) state that the sequence $\{w_k(\mathbf{0})\}$ is non-increasing and converges to a limit in K samples, and give an expression for both, the number of samples and the limit value $\overline{h}_K$. Item c) states that, when all measurements equal zero, the accuracy obtained is identical to the true, best possible accuracy among causal differentiators formulated in Problem 1. Item d) shows that the zero-measurement case is actually the worst over all possible measurements. This means that the proposed differentiator's worst-case accuracy is thus the best among all causal differentiators. Note that these results are very powerful because $\mathbf{x}_k \in \mathcal{C}_k(\mathbf{m}_k)$ does not imply that $\mathbf{x}_k = \mathcal{X}_k(\mathfrak{f})$ for some $\mathfrak{f}$ satisfying (1)–(2), as stated in Remark 1.

Since the best worst-case accuracy is achieved after a fixed number of K sampling steps and then stays constant, considering more (older) measurements does not improve the worst-case performance. With this insight, Theorem 1 ensures that Algorithm 1 with $\hat{K} \geq K$ provides the best worst-case accuracy among all causal differentiators. Particularly, if $N < LT^2/4$, then $\hat{K} = K = 1$ can be chosen; the linear programs may then be solved explicitly, yielding the differentiator $\hat{\mathfrak{f}}_k^1 = (m_k - m_{k-1})/T$ as a special case.

IV. PROOF OF THEOREM 1

The proof strategy for Theorem 1 is to first study the case corresponding to the noise bound $N = 1$ and the sampling period $T = 1$, and then show how the general case can be obtained from this.

A. The case $N = T = 1$

To begin, consider the case $(N, T, L) = (1, 1, \tilde{L})$ with $\tilde{L} := 4/\varepsilon^2$. Using these parameters, the quantities in the statement of Theorem 1 become

$$K = \begin{cases} \lfloor \varepsilon \rfloor & \text{if } \lfloor \varepsilon \rfloor^2 + \lfloor \varepsilon \rfloor \geq \varepsilon^2, \\ \lfloor \varepsilon \rfloor + 1 & \text{otherwise}, \end{cases} \tag{18}$$

$$\overline{a}_k = h(\underline{k}), \quad \underline{k} = \min\{k, K\}, \quad h(\ell) = \frac{2\ell}{\varepsilon^2} + \frac{2}{\ell},$$

where we have used $\overline{a}_k$ and $h(\cdot)$ to denote $\overline{h}_k$ and $h_o(\cdot)$ corresponding to $(N, T, L) = (1, 1, 4/\varepsilon^2)$.

The following lemma establishes some properties of the sequence $\{\overline{a}_k\}$ that will be required next.

Lemma 1: Consider (18), and let

$$a_k := \min_{\ell \in \{1, \ldots, k\}} h(\ell).$$

Then, the following statements are true:
a) $h(\ell)$ is strictly decreasing for $\ell \leq \lfloor \varepsilon \rfloor$ and strictly increasing for $\ell \geq \lfloor \varepsilon \rfloor + 1$.
b) If $k \leq \lfloor \varepsilon \rfloor$ then $a_k = h(k)$.
c) If $k \geq \lfloor \varepsilon \rfloor + 1$ then $a_k = a_K$.
d) $\overline{a}_k = a_k$ for all $k \in \mathbb{N}$.

Proof: The derivative of the function h is $h^{(1)}(s) = \frac{2}{\varepsilon^2} - \frac{2}{s^2}$, so that $h^{(1)}(\varepsilon) = 0$, $h^{(1)}(s) < 0$ for $s \in (0, \varepsilon)$, and

$h^{(1)}(s) > 0$ for $s \in (\varepsilon, \infty)$. Therefore, h is strictly decreasing within the interval $(0, \varepsilon]$ and strictly increasing within $[\varepsilon, \infty)$. Since $\lfloor \varepsilon \rfloor \leq \varepsilon < \lfloor \varepsilon \rfloor + 1$, then item a) is established. Item b) then follows straightforwardly from the definition of a_k.

For item c), note that for $k \geq \lfloor \varepsilon \rfloor + 1 > \varepsilon$, from item a) we must have $a_k = \min\{h(\lfloor \varepsilon \rfloor), h(\lfloor \varepsilon \rfloor + 1)\}$. Consider

$$h(\lfloor \varepsilon \rfloor) - h(\lfloor \varepsilon \rfloor + 1) = \frac{2}{\lfloor \varepsilon \rfloor} - \frac{2}{\lfloor \varepsilon \rfloor + 1} - \frac{2}{\varepsilon^2}$$

If this difference is nonpositive, which happens if $\lfloor \varepsilon \rfloor^2 + \lfloor \varepsilon \rfloor \geq \varepsilon^2$, then $h(\lfloor \varepsilon \rfloor + 1) \geq h(\lfloor \varepsilon \rfloor)$ will hold. Observing (18), then $a_k = \min\{h(\lfloor \varepsilon \rfloor), h(\lfloor \varepsilon \rfloor + 1)\} = a_K$.

Finally, d) follows by combining Lemma 1b) and c). ∎

Let $S_k := \{\ell \in \mathbb{N} : 1 \leq \ell \leq k - \underline{k}\}$. Consider a function $\tilde{\mathfrak{f}} : [0, k] \to \mathbb{R}$ defined as follows

$$\tilde{\mathfrak{f}}(t) := \frac{2(t-k)^2}{\varepsilon^2} + \bar{a}_k(t-k) + 1 \text{ for } t \in [k - \underline{k}, k] \quad (19)$$

$$\tilde{\mathfrak{f}}(t) := \tilde{\mathfrak{f}}^{(1)}(\ell)(t-\ell)^2 + \tilde{\mathfrak{f}}^{(1)}(\ell)(t-\ell) - 1 \quad (20)$$

for $t \in [\ell - 1, \ell)$ with $\ell \in S_k$ and where

$$\tilde{\mathfrak{f}}^{(1)}(\ell) := \lim_{t \to \ell^+} \tilde{\mathfrak{f}}^{(1)}(t). \quad (21)$$

It is clear that $\tilde{\mathfrak{f}}$ satisfies $\lim_{t \to k^-} \tilde{\mathfrak{f}}^{(1)}(t) = \bar{a}_k$. The following lemma establishes that $\tilde{\mathfrak{f}}$ is continuously differentiable, its derivative has a global Lipschitz constant $\tilde{L}$, and satisfies $\mathcal{X}_k(\tilde{\mathfrak{f}}) \in \mathcal{C}_k(\mathbf{0})$.

Lemma 2: Let $\tilde{\mathfrak{f}} : [0, k] \to \mathbb{R}$ be defined by (19)–(21). Then,
a) $\tilde{\mathfrak{f}}$ is continuous, $\tilde{\mathfrak{f}}(k) = 1$, and $\tilde{\mathfrak{f}}(\ell) = -1$ for $\ell \in S_k$.
b) $\tilde{\mathfrak{f}}^{(1)}$ is continuous in $(0, k)$, and $\tilde{\mathfrak{f}}^{(1)}(\ell - 1) = -\tilde{\mathfrak{f}}^{(1)}(\ell)$ for every $\ell \in S_k$.
c) $|\tilde{\mathfrak{f}}^{(1)}(k - \underline{k})| \leq 2/\varepsilon^2$ for $k > \underline{k}$
d) $\tilde{\mathfrak{f}}(\ell) \in [-1, 1]$ for every $\ell \in \mathbb{N}_0$ with $\ell \leq k$.
e) $|\tilde{\mathfrak{f}}^{(2)}(t)| \leq \tilde{L} = 4/\varepsilon^2$ for almost every $t \in [0, k]$.

Proof: a) The fact that $\tilde{\mathfrak{f}}(k) = 1$ follows directly from (19). Also, since $\bar{a}_k = h(\underline{k})$, then

$$\tilde{\mathfrak{f}}(k - \underline{k}) = \frac{2\underline{k}^2}{\varepsilon^2} + \left(\frac{2\underline{k}}{\varepsilon^2} + \frac{2}{\underline{k}} \right)(-\underline{k}) + 1 = -1.$$

Note that, by definition, $\tilde{\mathfrak{f}}$ is continuous in the intervals $[\ell - 1, \ell)$ for all $\ell \in S_k$ and also in $[k - \underline{k}, k]$. From (20), for $\ell \in S_k$ one has $\lim_{t \to \ell^-} \tilde{\mathfrak{f}}(t) = -1 = \tilde{\mathfrak{f}}(\ell)$. Thus, $\tilde{\mathfrak{f}}(\ell) = -1$ for any $\ell \in S_k$ and it follows that $\tilde{\mathfrak{f}}$ is continuous in $[0, k]$.

b) From (20) we obtain:

$$\tilde{\mathfrak{f}}^{(1)}(t) = 2\tilde{\mathfrak{f}}^{(1)}(\ell)(t-\ell) + \tilde{\mathfrak{f}}^{(1)}(\ell), \quad t \in [\ell - 1, \ell) \quad (22)$$

Hence, (22) gives $\lim_{t \to \ell^-} \tilde{\mathfrak{f}}^{(1)}(t) = \tilde{\mathfrak{f}}^{(1)}(\ell)$ which according to (21) leads to continuity of $\tilde{\mathfrak{f}}^{(1)}(t)$ for every $t \in S_k$. Continuity of $\tilde{\mathfrak{f}}^{(1)}$ within $(0, k]$ then follows similarly as in the proof of item a. Finally, note that evaluating at $t = \ell - 1$ in (22) it follows that $\tilde{\mathfrak{f}}^{(1)}(\ell - 1) = -\tilde{\mathfrak{f}}^{(1)}(\ell)$ for every $\ell \in S_k$.

c) From (18), (19) and the definition $\bar{a}_k = h(\underline{k})$, it follows that for $k > \underline{k} = K$

$$\tilde{\mathfrak{f}}^{(1)}(k - \underline{k}) = \frac{2}{K} - \frac{2K}{\varepsilon^2}. \quad (23)$$

Multiplying $\tilde{\mathfrak{f}}^{(1)}(k - \underline{k})$ by $K\varepsilon^2/2 > 0$, the inequalities $-K \leq \varepsilon^2 - K^2 \leq K$ have to be proven. Consider first the case

$K = \lfloor \varepsilon \rfloor$. Then, $K \leq \varepsilon$, i.e., $\varepsilon^2 - K^2 \geq 0$, and the upper bound remains to be proven. Since $\lfloor \varepsilon \rfloor^2 + \lfloor \varepsilon \rfloor \geq \varepsilon^2$ holds in this case, one has

$$\varepsilon^2 - \lfloor \varepsilon \rfloor^2 \leq \varepsilon^2 - (\varepsilon^2 - \lfloor \varepsilon \rfloor) = \lfloor \varepsilon \rfloor = K.$$

Consider now the case $K = \lfloor \varepsilon \rfloor + 1$. Since $K > \varepsilon$, it suffices to show the lower bound. It is obtained from

$$\varepsilon^2 - (\lfloor \varepsilon \rfloor + 1)^2 = \varepsilon^2 - \lfloor \varepsilon \rfloor^2 - 2\lfloor \varepsilon \rfloor - 1 > -\lfloor \varepsilon \rfloor - 1 = -K,$$

because $\lfloor \varepsilon \rfloor^2 + \lfloor \varepsilon \rfloor < \varepsilon^2$ holds in this case.

d) For $\ell \in S_k$, item d) follows from item a). Otherwise, for $\ell \geq k - \underline{k} + 1$, obtain from the time derivative of (19) for $t \geq k - \underline{k} + 1$

$$\tilde{\mathfrak{f}}^{(1)}(t) \geq \tilde{\mathfrak{f}}^{(1)}(k - \underline{k} + 1) = \frac{4}{\varepsilon^2} + \tilde{\mathfrak{f}}^{(1)}(k - \underline{k}) \geq \frac{2}{\varepsilon^2} > 0 \quad (24)$$

for $k > \underline{k}$ due to item c) and

$$\tilde{\mathfrak{f}}^{(1)}(t) \geq \tilde{\mathfrak{f}}^{(1)}(1) = -\frac{4(k-1)}{\varepsilon^2} + \bar{a}_k = -\frac{2k-4}{\varepsilon^2} + \frac{2}{k}$$

$$\geq -\frac{2\varepsilon - 2}{\varepsilon^2} + \frac{2}{\varepsilon + 1} = \frac{2}{\varepsilon^2(\varepsilon + 1)^2} > 0 \quad (25)$$

for $k = \underline{k} \leq K \leq \varepsilon + 1$. Hence, $\tilde{\mathfrak{f}}$ is strictly increasing on $[k - \underline{k} + 1, k]$ and, since $\tilde{\mathfrak{f}}(k) = 1$, it suffices to show $\tilde{\mathfrak{f}}(k - \underline{k} + 1) \geq -1$. To see this, assume the opposite

$$-1 > \tilde{\mathfrak{f}}(k - \underline{k} + 1) = -1 + \frac{2}{\underline{k}} + \frac{2}{\varepsilon^2} - \frac{2\underline{k}}{\varepsilon^2}$$

or equivalently that $\varepsilon^2 + \underline{k} - \underline{k}^2 < 0$. For $\underline{k} \leq \lfloor \varepsilon \rfloor$ this is impossible, because then $\varepsilon \geq \underline{k}$; hence, $\underline{k} = \lfloor \varepsilon \rfloor + 1 = K$. Then, $\varepsilon^2 + \lfloor \varepsilon \rfloor + 1 - (\lfloor \varepsilon \rfloor + 1)^2 < 0$ or equivalently $\varepsilon^2 < \lfloor \varepsilon \rfloor^2 + \lfloor \varepsilon \rfloor$ which contradicts the fact that $K = \lfloor \varepsilon \rfloor + 1$.

e) Note that $\tilde{\mathfrak{f}}^{(2)}(t) = 4/\varepsilon^2$ for $t \in (k - \underline{k}, k)$. The result is thus established if $\underline{k} = k$. Next, consider $k > \underline{k} = K$. From (20), $|\tilde{\mathfrak{f}}^{(2)}(t)| = |2\tilde{\mathfrak{f}}^{(1)}(\ell)|$ for $t \in (\ell - 1, \ell)$ with $\ell \in S_k$. From item b), then $|2\tilde{\mathfrak{f}}^{(1)}(\ell)| = |2\tilde{\mathfrak{f}}^{(1)}(k - \underline{k})|$ for $\ell \in S_k$, where $|2\tilde{\mathfrak{f}}^{(1)}(k - \underline{k})| \leq 4/\varepsilon^2 = \tilde{L}$ due to item c). Thus, $|\tilde{\mathfrak{f}}^{(2)}(t)| \leq \tilde{L}$ follows for almost every $t \in [0, k]$. ∎

B. The case with arbitrary positive N, T, L

Let $\mathfrak{f}(t) := N\tilde{\mathfrak{f}}(t/T)$ for $t \in [0, kT]$ and $\tilde{\mathfrak{f}}(t)$ defined as in (19) and (20) with $\varepsilon = \frac{2}{T}\sqrt{\frac{N}{L}}$. First, Lemma 2, items a) and b), is used to conclude that $\mathfrak{f}$ is continuously differentiable in $(0, kT)$. Next, Lemma 2d) is used to conclude that $\mathfrak{f}(\ell T) \in [-N, N]$ for every integer $\ell \in [0, k]$. Moreover, Lemma 2e) is used to conclude that for almost all $t \in [0, kT]$,

$$\left| \mathfrak{f}^{(2)}(t) \right| = \frac{N}{T^2} \left| \tilde{\mathfrak{f}}^{(2)}(t/T) \right| \leq \frac{N}{T^2} \frac{4}{\varepsilon^2} = \frac{N}{T^2} \frac{LT^2}{N} = L.$$

Furthermore, using $\varepsilon = (2/T)\sqrt{N/L}$ in (18) recovers the definitions in the statement of Theorem 1 directly.

It follows that the function $\mathfrak{f}$ satisfies (1)–(2) for some sequence $\{\eta_k\}$ and zero measurements, and hence $\mathcal{X}_k(\mathfrak{f}) \in \mathcal{C}_k(\mathbf{0})$ (recall Remark 1). In addition, $\mathfrak{f}_k^1 = \mathfrak{f}^{(1)}(kT) = \bar{h}_k = N\bar{a}_k/T$. From (5), (12) and (13), then

$$\overline{f}_k^1(\mathbf{0}) \geq \overline{\mathcal{F}}_k^1(\mathbf{0}) \geq \mathfrak{f}_k^1 = \bar{h}_k. \quad (26)$$

Next, $\bar{h}_k$ is shown to be also an upper bound for $\overline{f}_k^1(\mathbf{0})$.

Lemma 3: Let $k, \underline{k} \in \mathbb{N}$ satisfy $k \geq \underline{k}$. Consider real numbers f_j, f_j^1 for $j = 0, 1, \ldots, k$ satisfying in (11) the inequalities corresponding to (7) and (8) for $j = k - \underline{k} + 1, \ldots, k$, and to (9) for $j = k$ and $j = k - \underline{k}$. Let $\bar{h}_k$ be defined as in Theorem 1. Then, $f_k^1 \leq \bar{h}_k$.

Proof: From (8) we know that $f_{j-1}^1 \geq f_j^1 - LT$. Using this relation repeatedly for $j = k, k-1, \ldots, k - \underline{k} + 1$,

$$f_{k-i}^1 \geq f_k^1 - iLT \quad \text{for } i = 1, \ldots, \underline{k}. \tag{27}$$

Similarly, from (7), we know that $f_{j-1} \leq f_j - f_j^1 T + T^2 L/2$. Using this relation for $j = k, k-1, \ldots, k - \underline{k} + 1$ yields

$$f_{k-i} \leq f_k - \sum_{j=0}^{i-1} f_{k-j}^1 T + iL\frac{T^2}{2}, \quad i = 1, \ldots, \underline{k}. \tag{28}$$

Let $i = \underline{k}$ and use (27) in (28) to obtain:

$$f_{k-\underline{k}} \leq f_k - \sum_{j=0}^{\underline{k}-1} \left(f_k^1 - jLT\right) T + \underline{k}L\frac{T^2}{2}$$
$$\leq f_k - \underline{k}T f_k^1 + LT^2 \underline{k}^2/2. \tag{29}$$

Using $-LT^2\underline{k}^2/2 = 2N - \underline{k}T\bar{h}_k$ from the definition of $\bar{h}_k$,

$$f_k - f_{k-\underline{k}} \geq \underline{k}T f_k^1 - \frac{\underline{k}^2 LT^2}{2} = 2N + \underline{k}T(f_k^1 - \bar{h}_k). \tag{30}$$

However, from (9) we know that, $f_k \leq N$ and $-f_{k-\underline{k}} \leq N$. Thus, $f_k - f_{k-\underline{k}} \leq 2N$, which with (30) yields $f_k^1 \leq \bar{h}_k$. ∎

Combining (26) and Lemma 3 leads to $\overline{f}_k^1(\mathbf{0}) = \overline{\mathcal{F}}_k^1(\mathbf{0}) = \bar{h}_k$. From (11), it follows that $\mathbf{x} \in \mathcal{C}_k(\mathbf{0}) \Leftrightarrow -\mathbf{x} \in \mathcal{C}_k(\mathbf{0})$. Therefore, it must happen that $\underline{f}_k^1(\mathbf{0}) = \underline{\mathcal{F}}_k^1(\mathbf{0}) = -\bar{h}_k$. Finally, recalling (14) and (15), then $w_k(\mathbf{0}) = \mathcal{W}_k(\mathbf{0}) = 2\bar{h}_k$. This establishes Theorem 1a) and c).

To prove item b), note that $\bar{h}_k$ in Theorem 1 satisfies $\bar{h}_k = (N/T)h(\underline{k})$, with the latter defined as in (18) and $\varepsilon^2 = 4N/LT^2$. Therefore, Theorem 1b) follows from Lemma 1.

C. The case with $\mathbf{m}_k \neq \mathbf{0}$

The constraint set $\mathcal{C}_k(\mathbf{m}_k)$ has the following simple property, which will be instrumental in establishing Theorem 1d).

Lemma 4: Let $\mathbf{m}_k \in \mathbb{R}^{k+1}$ as in (4) and $\mathbf{x}_k \in \mathbb{R}^{2k+2}$ with components named as in (10) be such that $\mathbf{x}_k \in \mathcal{C}_k(\mathbf{m}_k)$. Let $\tilde{\mathbf{m}}_k \in \mathbb{R}^{k+1}$ have components $\tilde{m}_0, \tilde{m}_1, \ldots, \tilde{m}_k$ satisfying

$$\tilde{m}_j = m_j + aj + b, \quad j = 0, 1, \ldots, k, \tag{31}$$

for some $a, b \in \mathbb{R}$ and define $\tilde{\mathbf{x}}_k := [(\tilde{f}_{0:k})^T, (\tilde{f}_{0:k}^1)^T]^T$, with

$$\tilde{f}_j := f_j + aj + b, \quad \tilde{f}_j^1 := f_j^1 + a/T, \quad j = 0, 1, \ldots, k.$$

Then, $\tilde{\mathbf{x}}_k \in \mathcal{C}_k(\tilde{\mathbf{m}}_k)$.

Proof: Directly from the definitions, it is clear that

$$\tilde{f}_j - \tilde{m}_j = f_j - m_j, \quad \tilde{f}_j^1 - \tilde{f}_{j-1}^1 = f_j^1 - f_{j-1}^1,$$
$$\tilde{f}_{j-1} - \tilde{f}_j + \tilde{f}_j^1 T = f_{j-1} - f_j + f_j^1 T.$$

Therefore, if (7)–(9) are satisfied for $\mathbf{x}_k$ and $\mathbf{m}_k$, they will also be satisfied for $\tilde{\mathbf{x}}_k$ and $\tilde{\mathbf{m}}_k$. ∎

Consider $\mathbf{m}_k$ with nonempty $\mathcal{C}_k(\mathbf{m}_k)$ and let $\tilde{\mathbf{m}}_k$ be defined as in (31) with $a = (m_{k-\underline{k}} - m_k)/\underline{k}$ and $b = -m_k - ak$.

Then, $\tilde{m}_{k-\underline{k}} = \tilde{m}_k = 0$. By Lemma 4, it follows that $\overline{\tilde{f}}_k^1(\tilde{\mathbf{m}}_k) = \overline{f}_k^1(\mathbf{m}_k) + a/T$ and $\underline{\tilde{f}}_k^1(\tilde{\mathbf{m}}_k) = \underline{f}_k^1(\mathbf{m}_k) + a/T$, so that $w_k(\tilde{\mathbf{m}}_k) = w_k(\mathbf{m}_k)$.

Next, apply Lemma 3 to $\tilde{\mathbf{x}}_k \in \mathcal{C}_k(\tilde{\mathbf{m}}_k)$. This gives $\overline{\tilde{f}}_k^1(\tilde{\mathbf{m}}_k) \leq \bar{h}_k$. By the symmetry of the constraints required by Lemma 3, also $\underline{\tilde{f}}_k^1(\tilde{\mathbf{m}}_k) \geq -\bar{h}_k$. Therefore,

$$w_k(\mathbf{0}) = 2\bar{h}_k \geq w_k(\tilde{\mathbf{m}}_k) = w_k(\mathbf{m}_k) \geq \mathcal{W}_k(\mathbf{m}_k)$$

for every k and $\mathbf{m}_k$. Taking the supremum over all $\mathbf{m}_k$ yields $w_k(\mathbf{0}) \geq \bar{w}_k(L, N, T) \geq \overline{\mathcal{W}}_k(L, N, T) \geq \mathcal{W}_k(\mathbf{0})$. Theorem 1d) is then established recalling Theorem 1c).

V. COMPARISONS

This section compares the proposed differentiator's performance and accuracy to a linear high-gain and an exact sliding-mode differentiator. For comparison purposes, each of those two differentiators is discretized using state-of-the-art techniques. The proposed differentiator in Algorithm 1 is implemented by solving the linear programs using Yalmip [19] with the Matlab solver `linprog`.

Before doing the comparison, it is worthwile to note that Theorem 1 states the proposed differentiator's worst-case accuracy $\frac{1}{2}\mathcal{W}_k(L, N, T)$ and the maximum time KT it takes to achieve it. For all values of T, L, and N, this accuracy is bounded from below by

$$\frac{1}{2}\bar{\mathcal{W}}_k(L, N, T) \geq 2\sqrt{NL}. \tag{32}$$

This lower limit is also obtained for certain special combinations of T, L, N, as well as for $T \to 0$. Exact differentiators have a similar inherent accuracy restriction, see [10], [20].

A. Linear High-Gain Differentiator

In continuous time, a second order linear (high-gain) differentiator with identical eigenvalues and time constant τ is given by

$$\dot{y}_1 = \frac{2}{\tau}(m - y_1) + y_2, \qquad \dot{y}_2 = \frac{1}{\tau^2}(m - y_1), \tag{33}$$

with output $\hat{f}^1 = y_2$, input $m = f + \eta$ and $|\eta| \leq N$. From [7], its optimal asymptotic accuracy is obtained as $4e^{-\frac{1}{2}}\sqrt{NL} \approx 2.43\sqrt{NL}$. The corresponding optimal time constant is $\tau = e^{-\frac{1}{2}}\sqrt{N/L}$, which is hence chosen in the following. For simulation purposes, the linear system (33) is discretized using the implicit Euler method.

B. Robust Exact Sliding-Mode Differentiator

As a sliding-mode differentiator, the robust exact differentiator proposed in [10] is used. In continuous time, it is

$$\dot{y}_1 = k_1 |m - y_1|^{\frac{1}{2}} \operatorname{sign}(m - y_1) + y_2, \tag{34a}$$
$$\dot{y}_2 = k_2 \operatorname{sign}(m - y_1), \tag{34b}$$

with output $\hat{f}^1 = y_2$, input $m = f + \eta$ with $|\eta| \leq N$ and positive parameters k_1, k_2. It is discretized using the matching method proposed in [16] and simulated using the toolbox [21]. Parameters are selected as $k_1 = 2r$ and $k_2 = r^2$, with robustness factor r as in [21] set to $r = 1.5\sqrt{L}$.
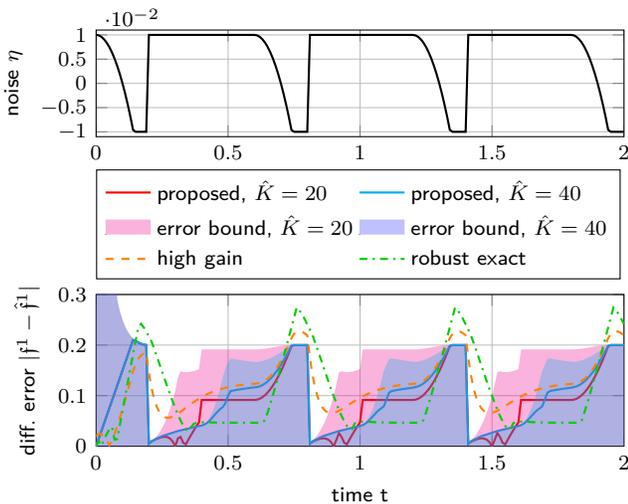
Fig. 1. Bounded noise η added to the signal $\mathfrak{f}(t) = t^2/2$ sampled with $T = 10^{-2}$, and corresponding differentiation errors for proposed differentiator, linear high-gain differentiator, and sliding-mode differentiator from a simulation with $N = 10^{-2}$ and $L = 1$. For the proposed differentiator, the error bounds obtained along with the estimate from the linear programs in Algorithm 1 are also shown.

C. Comparison

For the comparison, the signal $\mathfrak{f}(t) = Lt^2/2$ and noise

$$\eta(t) = \begin{cases} \max(-N, N - L(t - c\lfloor \frac{t}{c} \rfloor)^2) & t - c\lfloor \frac{t}{c} \rfloor < 2\sqrt{\frac{N}{L}} \\ N & \text{otherwise} \end{cases}$$

with constant $c = 6\sqrt{N/L}$ are sampled with $T = 10^{-2}$. Parameters are selected as $L = 1$, $N = 10^{-2}$. For these particular parameters, Theorem 1 yields $K = 20$ and an optimal worst-case accuracy $h_o(K) = 0.2$.

Fig. 1 depicts the noise as well as the differentiation error of all differentiators. For the proposed differentiator, two values of $\hat{K}$ are considered and the error bounds, i.e., the values of $(\overline{f}_k^1 - \underline{f}_k^1)/2$, as obtained from the linear program are shown as well. One can see that, after an initial transient of duration $KT = 0.2$, the proposed differentiator achieves the best worst-case accuracy of $h_o(K) = 0.2$, as expected from the theoretical results. Moreover, increasing $\hat{K}$ improves the error bound obtained along with the estimate. The high-gain differentiator leads to a larger but smoother error overall. The robust exact differentiator, finally, exhibits the largest worst-case errors, because it attempts to differentiate exactly also the noise, but is the most accurate one for constant noise.

VI. CONCLUSION

A differentiator for sampled signals based on linear programming was proposed. It is shown that the best worst-case accuracy is obtained with a fixed number of discrete-time measurements, which allows limiting its computational complexity. Comparisons to a linear high-gain differentiator and a standard sliding-mode differentiator exhibited a higher accuracy. However, depending on the sampling time, the increased accuracy comes at a higher computational cost.

APPENDIX

To show that, as stated in Remark 1, a nonempty constraint set $\mathcal{C}_k$ does not necessarily imply existence of a function $\mathfrak{f}$, consider $L = 2$, $T = 1$, $N = 0$, $k = 2$ and measurements $\mathbf{m} = [0, 0, 4]^T$. It is easy to check that every $\mathbf{x} \in \mathcal{C}_k(\mathbf{m})$ has the form $\mathbf{x} = [0, 0, 4, f_0^1, 1, 3]^T$ with $f_0^1 \in [-1, 3]$. By symmetry with respect to time reversal, any function $\mathfrak{f}$ satisfying (1) also has to satisfy $|\mathfrak{f}_j - \mathfrak{f}_{j+1} + \mathfrak{f}_j^1 T| \leq L\frac{T^2}{2}$ in addition to (7). Adding this inequality with $j = 1$ to the constraints as $|f_1 - f_2 + f_1^1| \leq 1$ yields a contradiction. Hence, no function $\mathfrak{f}$ satisfying (1) exists for these measurements.

REFERENCES

[1] M. Reichhartinger, D. Efimov, and L. Fridman, "Special issue on differentiators," *Int. J. Control*, vol. 91, no. 9, pp. 1980–1982, 2018.

[2] M. R. Mojallizadeh, B. Brogliato, and V. Acary, "Discrete-time differentiators: design and comparative analysis." HAL Preprint ID: hal-02960923, 2021.

[3] Y. Shtessel, C. Edwards, L. Fridman, and A. Levant, "Observation and Identification via HOSM Observers," in *Sliding Mode Control and Observation*, pp. 251–290, Springer, 2014.

[4] H. Ríos, J. Davila, L. Fridman, and C. Edwards, "Fault detection and isolation for nonlinear systems via high-order-sliding-mode multiple-observer," *Int. J. Robust Nonlin.*, vol. 25, no. 16, pp. 2871–2893, 2015.

[5] D. Efimov, L. Fridman, T. Raïssi, A. Zolghadri, and R. Seydou, "Application of interval observers and HOSM differentiators for fault detection," in *IFAC Proceedings*, vol. 8, pp. 516–521, 2012.

[6] F. J. Bejarano and L. Fridman, "High order sliding mode observer for linear systems with unbounded unknown inputs," *Int. J. Control*, vol. 83, no. 9, pp. 1920–1929, 2010.

[7] L. K. Vasiljevic and H. K. Khalil, "Error bounds in differentiation of noisy signals by high-gain observers," *Syst. Control Lett.*, vol. 57, no. 10, pp. 856–862, 2008.

[8] M. Mboup, C. Join, and M. Fliess, "A revised look at numerical differentiation with an application to nonlinear feedback control," in *Mediterranean Conference on Control and Automation*, IEEE, 2007.

[9] A. Othmane, J. Rudolph, and H. Mounier, "Analysis of the parameter estimate error when algebraic differentiators are used in the presence of disturbances," in *21st IFAC World Congress*, pp. 572–577, 2020.

[10] A. Levant, "Robust Exact Differentiation via Sliding Mode Technique," *Automatica*, vol. 34, no. 3, pp. 379–384, 1998.

[11] A. Levant, "Higher-order sliding modes, differentiation and output-feedback control," *Int. J. Control*, vol. 76, no. 9-10, pp. 924–941, 2003.

[12] L. Fraguela, M. T. Angulo, J. A. Moreno, and L. Fridman, "Design of a prescribed convergence time uniform Robust Exact Observer in the presence of measurement noise," in *Conference on Decision and Control*, pp. 6615–6620, 2012.

[13] A. Polyakov, D. Efimov, and B. Brogliato, "Consistent discretization of finite-time and fixed-time stable systems," *SIAM Journal on Control and Optimization*, vol. 57, no. 1, pp. 78–103, 2019.

[14] A. Levant, "On fixed and finite time stability in sliding mode control," in *Conference on Decision and Control*, pp. 4260–4265, IEEE, 2013.

[15] J. E. Carvajal-Rubio, J. D. Sánchez-Torres, A. G. Loukianov, M. Defoort, and M. Djemai, "Implicit and Explicit Discrete-Time Realizations of Homogeneous Differentiators," *Int. J. Robust Nonlin.*, 2021.

[16] S. Koch and M. Reichhartinger, "Discrete-time equivalent homogeneous differentiators," in *15th International Workshop on Variable Structure Systems (VSS)*, pp. 354–359, IEEE, 2018.

[17] F. Mazenc and O. Bernard, "Interval observers for linear time-invariant systems with disturbances," *Automatica*, vol. 47, no. 1, pp. 140–147, 2011.

[18] M. Guerra, C. Vázquez, D. Efimov, G. Zheng, L. Freidovich, and W. Perruquetti, "Interval differentiators: On-line estimation of differentiation accuracy," in *European Control Conference*, pp. 1347–1352, IEEE, 2017.

[19] J. Löfberg, "Yalmip: A toolbox for modeling and optimization in matlab," in *IEEE Int Conf Robot Autom*, pp. 284–289, 2004.

[20] A. Levant, M. Livne, and X. Yu, "Sliding-mode-based differentiation and its application," in *20th IFAC World Congress*, pp. 1699–1704, 2017.

[21] B. Andritsch, M. Horn, S. Koch, H. Niederwieser, M. Wetzlinger, and M. Reichhartinger, "The robust exact differentiator toolbox revisited: Filtering and discretization features," in *IEEE International Conference on Mechatronics (ICM)*, 2021.