2022 American Control Conference
Atlanta, USA, June 8-10, 2022

# On the Accuracy of Numerical Differentiation Using High-Gain Observers and Adaptive Input Estimation

Shashank Verma*, Sneha Sanjeevini*, E. Dogan Sumer+, Anouck Girard*, and Dennis S. Bernstein*

*Abstract*— **The ability to control a system is often enhanced by feeding back the derivatives of sensor signals, such as estimates of velocity and acceleration when only position is measured. Within this context, signal differentiation must be performed causally, that is, using only current and past data and with minimal computational latency. This paper formulates causal differentiation as a sampled-data input-estimation problem, where the plant is a cascade of integrators. Adaptive input estimation based on retrospective-cost optimization is considered, where the innovations from the Kalman filter is used to drive the online adaptation. Using backward-difference differentiation (BDD) as a baseline comparison, high-gain observers (HGO) with bilinear discretization and retrospective cost input estimation (RCIE) are applied to harmonic signals under various noise levels for single and double differentiation. These methods are then applied to experimental position data of a small rover for estimating its velocity and acceleration. Neither method uses information about the noise statistics, and no analog or digital filtering is used for noise suppression.**

## I. INTRODUCTION

As exemplified by PID control [1], [2], it is often of great benefit to be able to use the derivative of a sensor measurement as part of a feedback control law. Unfortunately, differentiation corresponds to an unbounded operator, which has unbounded gain; this property is evident in the fact that the derivative of small-amplitude high-frequency noise may have arbitrarily large amplitude. In practice, differentiation is typically performed by combining low-frequency approximate differentiation with lowpass filtering for noise suppression. The required shape of the filter is typically determined manually based on the characteristics of the sensor noise.

Within a digital context, numerical differentiation can be viewed as an approximation of "analog" differentiation, where the goal is to obtain a discrete-time signal that approximates the "analog" derivative for the given sample rate. A constraint on discrete-time differentiation for control-system applications is the fact that only data up to the present time can be used to compute the derivative; we call this *causal numerical differentiation*. For example, backward-difference differentiation is causal, whereas central-difference and forward-difference differentiation are not. Causal differentiation avoids the need for future data, which entails latency and thus may degrade performance and potentially lead to instability.

*Shashank Verma, Sneha Sanjeevini, Anouck Girard, and Dennis S. Bernstein are with the Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109, USA shaaero@umich.edu

+E. S. Dogan is with Ford Motor Company.

Given the importance of causal numerical differentiation, it is not surprising that substantial effort has been devoted to this problem. One of the earliest approaches to numerical differentiation is given in [3], where the problem is posed in terms of an integral equation, whose solution entails regularization. An extension of this approach based on total-variation regularization is given in [4], and numerical comparisons are given in [5]. This approach uses data from the entire interval and thus is not suitable for causal numerical differentiation. Motivated by the goal of developing alternatives to traditional observer design, an interpolation approach with a moving window is used in [6], [7], where the derivatives of the interpolating basis functions are used to approximate the derivative of the given function. The accuracy of this approach was shown to improve as the data window size increases, albeit at a cost of latency.

In order to estimate velocity and acceleration from camera data, ten numerical differentiation algorithms are compared in [8]. Although many of these algorithms performed well on experimental data, most of the techniques are based on a window involving both past and future data and thus entail latency when used for online applications.

In [9], an $n$th-order derivative is viewed as the input to a linear system whose output is the signal data. A high-gain observer is used to approximate the derivative of the output. To enable discrete-time implementation, the observer is then discretized using various methods including zero-order hold and bilinear transformation. Motivated by nonlinear feedback control, iterated integrals are used to represent lowpass filters in [10] and [11]. This technique enables computation of derivatives of arbitrary order with minimal latency.

An alternative approach to causal numerical differentiation is based on sliding-mode control. The extensive literature on this approach is reviewed in [12] with connections to the robust exact differentiator and the super-twisting algorithm. Online parameter adaptation is considered in [12] to provide adjustable tuning as the characteristics of the signal change. Further extensions of this approach are developed in [13]. The approach in [14] uses multi-objective optimization to select parameters that minimize the loss function to maintain the faithfulness and smoothness of the estimate.

The present paper considers the problem of causal numerical differentiation as an input-estimation problem. In input estimation, the input to a linear system is assumed to be unknown, and the measured output of the system is used to estimate the input of the system [15]–[28]. When the dynamics of the system consist of a cascade of one or more integrators, the estimates of the input provide estimates

978-1-6654-5197-0/$31.00 ©2022 AACC

of one or more derivatives of the output signal. Since, like state estimation, input estimation is an online technique, this approach is suitable for causal numerical differentiation. However, the works in [15]–[28] do not consider adaptive input estimation. Adaptive input estimation is considered in [29], where the input estimation is combined with state estimation based on the discrete-time Kalman filter. The adaptive input estimator uses retrospective cost optimization to provide an estimate of the input for use by the Kalman filter. The error metric for adaptation is given by the estimation residual, that is, the innovations.

The goal of the present paper is to compare the accuracy of retrospective cost input estimation (RCIE) given in [29] and the high-gain observer (HGO) given in [9] for causal numerical differentiation in the presence of noisy measurements. For simplicity and clarity, the numerical study in this paper considers harmonic signals corrupted by Gaussian white noise. For each input signal, backward-difference numerical differentiation provides a baseline for performance comparison. The present paper also applies the different methods to the position data of a small rover to estimate its velocity and acceleration.

## II. RETROSPECTIVE COST INPUT ESTIMATION

Retrospective cost input estimation (RCIE) [29] can be applied to MIMO linear time-varying systems. The objective in this paper is to use RCIE for differentiation, and for the purpose of differentiation, it is enough to consider SISO linear-time invariant systems. Hence, the RCIE algorithm is summarized here within the context of SISO LTI systems.

Consider the linear discrete-time system

$$x_{k+1} = Ax_k + Bd_k + D_1 w_k, \tag{1}$$

$$y_k = Cx_k + D_2 v_k, \tag{2}$$

where $k$ is step, $x_k \in \mathbb{R}^{l_x}$ is the unknown state, $d_k \in \mathbb{R}$ is the unknown input, $w_k \in \mathbb{R}$ is standard Guassian white process noise, $y_k \in \mathbb{R}$ is the measured output, and $v_k \in \mathbb{R}$ is standard Guassian white measurement noise. The matrices $A \in \mathbb{R}^{l_x \times l_x}$, $B \in \mathbb{R}^{l_x \times 1}$, $D_1 \in \mathbb{R}^{l_x \times 1}$, $C \in \mathbb{R}^{1 \times l_x}$, and $D_2 \in \mathbb{R}$ are assumed to be known. Define $V_1 \triangleq D_1 D_1^{\mathrm{T}}$ and $V_2 \triangleq D_2 D_2^{\mathrm{T}}$. The goal is to estimate $d_k$ and $x_k$.

### A. Input Estimation

Consider the Kalman filter forecast step

$$x_{\mathrm{fc},k+1} = Ax_{\mathrm{da},k} + B\hat{d}_k, \tag{3}$$

$$y_{\mathrm{fc},k} = Cx_{\mathrm{fc},k}, \tag{4}$$

$$z_k = y_{\mathrm{fc},k} - y_k, \tag{5}$$

where $\hat{d}_k$ is the estimate of $d_k$, $x_{\mathrm{da,k}} \in \mathbb{R}^{l_x}$ is the data-assimilation state, $x_{\mathrm{fc},k} \in \mathbb{R}^{l_x}$ is the forecast state, and $z_k \in \mathbb{R}$ is the innovations. The *input-estimation subsystem* of order $n_{\mathrm{c}}$ given by

$$\hat{d}_k = \sum_{i=1}^{n_{\mathrm{c}}} P_{i,k}\hat{d}_{k-i} + \sum_{i=0}^{n_{\mathrm{c}}} Q_{i,k} z_{k-i}, \tag{6}$$

where $P_{i,k} \in \mathbb{R}$ and $Q_{i,k} \in \mathbb{R}$, is used to obtain $\hat{d}_k$. RCIE minimizes $z_{k+1}$ by updating $P_{i,k}$ and $Q_{i,k}$. The subsystem (6) can be reformulated as

$$\hat{d}_k = \Phi_k \theta_k, \tag{7}$$

where the regressor matrix $\Phi_k$ is defined by

$$\Phi_k \triangleq \begin{bmatrix} \hat{d}_{k-1} & \cdots & \hat{d}_{k-n_{\mathrm{c}}} & z_k & \cdots & z_{k-n_{\mathrm{c}}} \end{bmatrix} \in \mathbb{R}^{1 \times l_\theta}, \tag{8}$$

the coefficient vector $\theta_k$ is defined by

$$\theta_k \triangleq \begin{bmatrix} P_{1,k} & \cdots & P_{n_{\mathrm{c}},k} & Q_{0,k} & \cdots & Q_{n_{\mathrm{c}},k} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{l_\theta}, \tag{9}$$

and $l_\theta \triangleq 2n_{\mathrm{c}} + 1$. In terms of the forward shift operator $\mathbf{q}$, (6) can be written as

$$\hat{d}_k = G_{\hat{d}z,k}(\mathbf{q}) z_k, \tag{10}$$

where

$$G_{\hat{d}z,k}(\mathbf{q}) \triangleq D_{\hat{d}z,k}^{-1}(\mathbf{q}) N_{\hat{d}z,k}(\mathbf{q}), \tag{11}$$

$$D_{\hat{d}z,k}(\mathbf{q}) \triangleq \mathbf{q}^{n_{\mathrm{c}}} - P_{1,k}\mathbf{q}^{n_{\mathrm{c}}-1} - \cdots - P_{n_{\mathrm{c}},k}, \tag{12}$$

$$N_{\hat{d}z,k}(\mathbf{q}) \triangleq Q_{0,k}\mathbf{q}^{n_{\mathrm{c}}} + Q_{1,k}\mathbf{q}^{n_{\mathrm{c}}-1} + \cdots + Q_{n_{\mathrm{c}},k}. \tag{13}$$

Next, define the filtered signals

$$\Phi_{\mathrm{f},k} \triangleq G_{\mathrm{f},k}(\mathbf{q})\Phi_k, \quad \hat{d}_{\mathrm{f},k} \triangleq G_{\mathrm{f},k}(\mathbf{q})\hat{d}_k. \tag{14}$$

Note that $G_{\mathrm{f},k}$ is a filter of window length $n_{\mathrm{f}} \geq 1$. Further details of the filter $G_{\mathrm{f},k}$ are given in the subsection II-C. Define the *retrospective performance variable*

$$z_{\mathrm{rc},k}(\hat{\theta}) \triangleq z_k - (\hat{d}_{\mathrm{f},k} - \Phi_{\mathrm{f},k}\hat{\theta}), \tag{15}$$

where the coefficient vector $\hat{\theta} \in \mathbb{R}^{l_\theta}$ denotes a variable for optimization, and define the retrospective cost function

$$J_k(\hat{\theta}) \triangleq \sum_{i=0}^{k} z_{\mathrm{rc},i}^{\mathrm{T}}(\hat{\theta}) R_z z_{\mathrm{rc},i}(\hat{\theta}) + \hat{\theta}^{\mathrm{T}}\Phi_i^{\mathrm{T}} R_d \Phi_i \hat{\theta} + (\hat{\theta} - \theta_0)^{\mathrm{T}} R_\theta (\hat{\theta} - \theta_0), \tag{16}$$

where $R_z \in \mathbb{R}$ is positive, $R_d \in \mathbb{R}$ is nonnegative, and $R_\theta \in \mathbb{R}^{l_\theta \times l_\theta}$ is positive definite. Define $P_0 \triangleq R_\theta^{-1}$. Then, for all $k \geq 0$, the cumulative cost function $J_k(\hat{\theta})$ has the unique global minimizer $\theta_{k+1}$ given by the RLS update

$$P_{k+1} = P_k - P_k \widetilde{\Phi}_k^{\mathrm{T}} \Gamma_k \widetilde{\Phi}_k P_k, \tag{17}$$

$$\theta_{k+1} = \theta_k - P_k \widetilde{\Phi}_k^{\mathrm{T}} \Gamma_k (\widetilde{z}_k + \widetilde{\Phi}_k \theta_k), \tag{18}$$

where

$$\Gamma_k \triangleq (\widetilde{R}^{-1} + \widetilde{\Phi}_k P_k \widetilde{\Phi}_k^{\mathrm{T}})^{-1}, \quad \widetilde{\Phi}_k \triangleq \begin{bmatrix} \Phi_{\mathrm{f},k} \\ \Phi_k \end{bmatrix}, \tag{19}$$

$$\widetilde{z}_k \triangleq \begin{bmatrix} z_k - \hat{d}_{\mathrm{f},k} \\ 0 \end{bmatrix}, \quad \widetilde{R} \triangleq \begin{bmatrix} R_z & 0 \\ 0 & R_d \end{bmatrix}. \tag{20}$$

Using the updated coefficient vector given by (18), the estimated input at step $k+1$ is given by replacing $k$ by $k+1$ in (7). We choose $\theta_0 = 0$, and thus $\hat{d}_0 = 0$.

### B. State Estimation

In order to estimate the state $x_k$, $x_{\mathrm{fc},k}$ given by (3) is used to obtain the estimate $x_{\mathrm{da},k}$ of $x_k$ given by the Kalman filter data-assimilation step

$$x_{\mathrm{da},k} = x_{\mathrm{fc},k} + K_{\mathrm{da},k} z_k, \tag{21}$$

where the state estimator gain $K_{\mathrm{da},k} \in \mathbb{R}^{l_x}$ is given by

$$K_{\mathrm{da},k} = -P_{\mathrm{f},k} C^{\mathrm{T}} (C P_{\mathrm{f},k} C^{\mathrm{T}} + V_2)^{-1}, \tag{22}$$

the data-assimilation error covariance $P_{\mathrm{da},k} \in \mathbb{R}^{l_x \times l_x}$ and the forecast error covariance $P_{\mathrm{f},k} \in \mathbb{R}^{l_x \times l_x}$ are given by

$$P_{\mathrm{da},k} = (I + K_{\mathrm{da},k} C) P_{\mathrm{f},k}, \tag{23}$$

$$P_{\mathrm{f},k+1} = A P_{\mathrm{da},k} A^{\mathrm{T}} + V_1 + \widetilde{V}_k, \tag{24}$$

where $\widetilde{V}_k \triangleq B$ var $(d_k - \hat{d}_k) B^{\mathrm{T}} + A$ cov $(x_k - x_{\mathrm{da},k}, d_k - \hat{d}_k) B^{\mathrm{T}} + B$ cov $(x_k - x_{\mathrm{da},k}, d_k - \hat{d}_k) A^{\mathrm{T}}$, and $P_{\mathrm{f},0} = 0$.

### C. The Filter $G_{\mathrm{f}}$

We choose $G_{\mathrm{f},k}(\mathbf{q})$ to be the FIR filter

$$G_{\mathrm{f},k}(\mathbf{q}) = \sum_{i=1}^{n_{\mathrm{f}}} H_{i,k} \frac{1}{\mathbf{q}^i}, \tag{25}$$

where, for all $k \geq 0$,

$$H_{i,k} \triangleq \begin{cases} CB, & k \geq i = 1, \\ C \left( \prod_{j=1}^{i-1} \overline{A}_{k-j} \right) B, & k \geq i \geq 2, \\ 0, & i > k, \end{cases} \tag{26}$$

and $\overline{A}_k \triangleq A(I + K_{\mathrm{da},k} C)$.

## III. CAUSAL NUMERICAL DIFFERENTIATION

In this section, the methods of performing causal[1] numerical differentiation using RCIE, high gain observer (HGO) and backward-difference differentiator (BDD) are explained.

### A. Differentiation using RCIE

Since the objective is to use RCIE as a differentiator, the system given by (1) and (2) is modeled as the discrete-time equivalent of an integrator. Thus, the measured output $y(t)$ is an integral of the unknown input $d(t)$ or, in other words, the unknown input $d(t)$ is the derivative of the measured output $y(t)$. Hence, by applying RCIE and reconstructing $\hat{d}$ from the estimates $\hat{d}_k$, we are estimating the derivative of the measured output $y$. Note that the concept of process noise is not applicable when the system is modeled as an integrator. Hence, for the rest of this paper, it is assumed that $w = 0$, and thus $D_1 = 0$.

Consider the $n$-th order integrator dynamics

$$\dot{x} = A_{\mathrm{I}} x + B_{\mathrm{I}} y^{(n)}, \quad y = C_{\mathrm{I}} x, \tag{27}$$

$$A_{\mathrm{I}} \triangleq \begin{bmatrix} 0_{(n-1)\times 1} & I_{n-1} \\ 0 & 0_{1\times(n-1)} \end{bmatrix}, \quad B_{\mathrm{I}} \triangleq \begin{bmatrix} 0_{(n-1)\times 1} \\ 1 \end{bmatrix}, \tag{28}$$

$$C_{\mathrm{I}} \triangleq \begin{bmatrix} 1 & 0_{1\times(n-1)} \end{bmatrix}, \tag{29}$$

where $x, y \in \mathbb{R}$, and $y^{(n)}$ is the $n$-th derivative of $y$. The discretization of (27) using zero-hold results in the discrete-time state space model given by

$$x_{k+1} = A_{\mathrm{d}} x_k + B_{\mathrm{d}} y_k^{(n)}, \quad y_k = C_{\mathrm{I}} x_k, \tag{30}$$

$$A_{\mathrm{d}} \triangleq e^{A_{\mathrm{I}} T_{\mathrm{s}}}, \quad B_{\mathrm{d}} \triangleq \int_0^{T_{\mathrm{s}}} e^{A_{\mathrm{I}}(t-\tau)} B_{\mathrm{I}} d\tau, \tag{31}$$

where $x_k \triangleq x(kT_{\mathrm{s}})$, $y_k \triangleq y(kT_{\mathrm{s}})$, $y_k^{(n)} \triangleq y^{(n)}(kT_{\mathrm{s}})$, and $T_{\mathrm{s}}$ is the sampling time. Setting $A = A_{\mathrm{d}}$, $B = B_{\mathrm{d}}$, and $C = C_{\mathrm{I}}$ in (1) and (2), and applying RCIE gives an estimate $(\hat{y}^{(n)} = \hat{d})$ of $y^{(n)}$. Note that $A_{\mathrm{d}} = 1, B_{\mathrm{d}} = T_{\mathrm{s}}$, and $C_{\mathrm{I}} = 1$ in the case of single differentiation, and

$$A_{\mathrm{d}} = \begin{bmatrix} 1 & T_{\mathrm{s}} \\ 0 & 1 \end{bmatrix}, \quad B_{\mathrm{d}} = \begin{bmatrix} \frac{1}{2} T_{\mathrm{s}}^2 \\ T_{\mathrm{s}} \end{bmatrix}, \quad C_{\mathrm{I}} = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{32}$$

in the case of double differentiation.

### B. Differentiation using HGO

The discrete-time implementation of casual differentiation using HGO was done in [9] and the same is explained here. A state space model for a high-gain observer designed for the system represented by (27) is given by

$$\dot{\hat{x}} = A_{\mathrm{co}} \hat{x} + B_{\mathrm{co}} y, \quad \hat{y} = C_{\mathrm{o}} \hat{x}, \tag{33}$$

$$A_{\mathrm{co}} \triangleq A_{\mathrm{I}} - H C_{\mathrm{I}}, \quad C_{\mathrm{o}} \triangleq \begin{bmatrix} 0_{(n-1)\times 1} & I_{n-1} \end{bmatrix}, \tag{34}$$

$$B_{\mathrm{co}} = H \triangleq \begin{bmatrix} \frac{\alpha_1}{\varepsilon} & \frac{\alpha_2}{\varepsilon^2} & \dots & \frac{\alpha_n}{\varepsilon^n} \end{bmatrix}^{\mathrm{T}}, \tag{35}$$

where $\varepsilon$ is a small positive parameter, and $\alpha_1, \alpha_2, \dots, \alpha_n$ are constants chosen such that the polynomial

$$p(s) \triangleq s^n + \alpha_1 s^{n-1} + \cdots + \alpha_{n-1} s + \alpha_n \tag{36}$$

is Hurwitz. The transfer function from $y$ to $\hat{y}$ is given by

$$G(s) = C_{\mathrm{o}} (sI - A_{\mathrm{I}} + H C_{\mathrm{I}})^{-1} H = D_G^{-1}(s) N_G(s), \tag{37}$$

where

$$D_G(s) \triangleq \varepsilon^n s^n + \alpha_1 \varepsilon^{n-1} s^{n-1} + \cdots + \alpha_{n-1} \varepsilon s + \alpha_n, \tag{38}$$

$$N_G(s) \triangleq \begin{bmatrix} \alpha_2 \varepsilon^{n-2} s^{n-1} + \cdots + \alpha_{n-1} \varepsilon s^2 + \alpha_n s \\ \alpha_3 \varepsilon^{n-3} s^{n-1} + \cdots + \alpha_{n-1} \varepsilon s^3 + \alpha_n s^2 \\ \vdots \\ \alpha_{n-1} \varepsilon s^{n-1} + \alpha_n s^{n-2} \\ \alpha_n s^{n-1} \end{bmatrix}. \tag{39}$$

Let

$$\hat{x} = \begin{bmatrix} \hat{x}_1 & \hat{x}_2 & \dots & \hat{x}_n \end{bmatrix}^{\mathrm{T}}, \tag{40}$$

$$\hat{y} = \begin{bmatrix} \hat{y}^{(1)} & \hat{y}^{(2)} & \dots & \hat{y}^{(n-1)} \end{bmatrix}^{\mathrm{T}}. \tag{41}$$

Since

$$\lim_{\varepsilon \to 0} G(s) = \begin{bmatrix} s & \cdots & s^{n-2} & s^{n-1} \end{bmatrix}^{\mathrm{T}}, \tag{42}$$

it follows that, for all $i = 1, \dots, n-1$, $\hat{y}^{(i)}$ is an approximation of $y^{(i)}$. Let the discrete-time observer state space model obtained by using bilinear transformation on (33) be

$$\hat{x}_{k+1} = A_{\mathrm{do}} \hat{x}_k + B_{\mathrm{do}} y_k, \quad \hat{y}_k = C_{\mathrm{o}} \hat{x}_k. \tag{43}$$

[1]The estimation of the derivative of $y_k$ uses the data $y_k$ and hence the estimation of the derivative of $y_k$ starts at step $k$. This implies that the estimate of derivative of $y_k$ is available only at step $k+1$ and thus there is a delay of one step in the estimation.

Thus, the implementation of (43) gives the estimates $(\hat{y}^{(1)}, \hat{y}^{(2)}, \ldots, \hat{y}^{(n-1)})$ of $y^{(1)}, y^{(2)}, \ldots, y^{(n-1)}$.

### C. Differentiation using BDD

The *single derivative backward-difference differentiator* is

$$G_{\rm sd}(\mathrm{z}) \triangleq \frac{\mathrm{z}-1}{T_{\rm s}\mathrm{z}}. \tag{44}$$

The *double derivative backward-difference differentiator* is

$$G_{\rm dd}(\mathrm{z}) \triangleq \frac{(\mathrm{z}-1)^2}{(T_{\rm s}\mathrm{z})^2}. \tag{45}$$

## IV. NUMERICAL EXAMPLES

In this section, numerical examples are given to illustrate the accuracy of RCIE and HGO as differentiators. BDD will be used as a baseline for comparison. Note that the examples will deal with discrete-time signals only.

*Example 4.1: Differentiation in the Absence of Noise.*

In this example, it is assumed that there is no output noise, and hence $v \equiv 0$. Let the measured output be $y_k = \sin(0.2k)$.

*Single Differentiation (SD):* In the case of RCIE, let $n_{\rm c} = 1$, $n_{\rm f} = 6$, $R_\theta = 10^{-3} I_3$, $R_d = 10^{-5}$, $R_z = 1$, $\widetilde{V} = 10^{-4}$. In the case of HGO, let $n = 3$ in (27), (28), and (29), let $\alpha_1 = 3, \alpha_2 = 3$, $\alpha_3 = 1$. Note that choosing $n = 3$ gave slightly better estimate of the first derivative as compared to choosing $n = 2$. The parameter $\varepsilon$ is chosen as the value between 0.01 and 2 that gives the lowest root mean square error (RMSE) between the estimated values and the true values. Figure 1 compares the signals estimated by SD/RCIE, SD/HGO, and

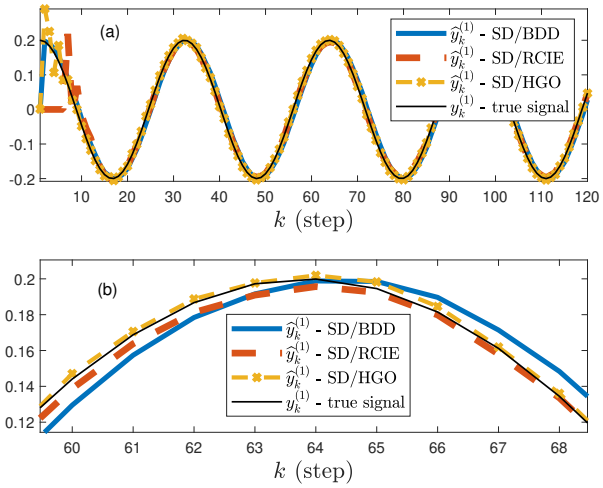


Fig. 1. *Example 4.1 Single Differentiation.* (a) The signals estimated by SD/RCIE and SD/HGO follow the true first derivative $y^{(1)}$ after about 20 steps, whereas the signal estimated by SD/BDD follows $y^{(1)}$ without a transient period. (b) A zoomed view of plot (a). At steady state, SD/HGO is more accurate than SD/RCIE and SD/BDD.

*Double Differentiation (DD):* In the case of RCIE, let $n_{\rm c} = 18$, $n_{\rm f} = 4$, $R_\theta = 10^{-1} I_{37}$, $R_d = 10^{-6}$, $R_z = 1$, $\widetilde{V} = 10^{-5}$. In the case of HGO, let $n = 4$ in (27), (28), and (29), let $\alpha_1 = 8, \alpha_2 = 24$, $\alpha_3 = 32$, $\alpha_4 = 16$. Note that choosing $n = 4$ gave slightly better estimate for the second derivative as compared to choosing $n = 3$. The parameter $\varepsilon$ is chosen in the same way as chosen for single differentiation. Figure

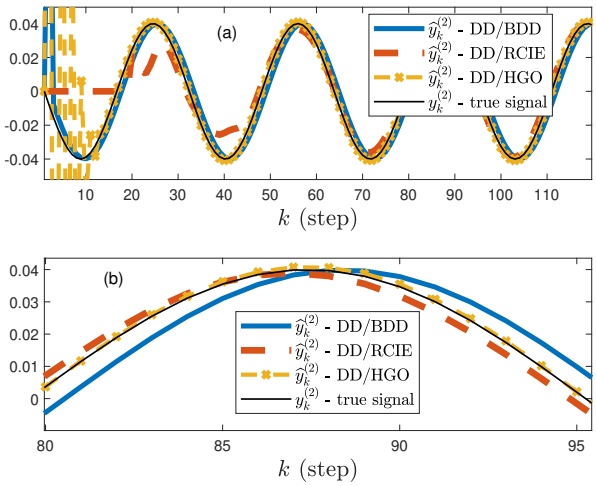


Fig. 2. *Example 4.1 Double Differentiation.* (a) The signal estimated by DD/HGO follows the true second derivative $y^{(2)}$ after about 20 steps, the signal estimated by DD/RCIE follows $y^{(2)}$ after about 50 steps, and the signal estimated by DD/BDD follows $y^{(2)}$ without a transient period. The signal estimated by DD/HGO has large oscillations in the transient period. (b) A zoomed view of plot (a). At steady state, DD/HGO is more accurate than DD/RCIE and DD/BDD.

2 compares the signals estimated by DD/RCIE, DD/HGO, and DD/BDD with the true second derivative.

*Example 4.2: Differentiation in the Presence of Noise.*

This example considers differentiation in the presence of output noise. Let the measured output be $y_k = \sin(0.2k) + D_2 v_k$, where $v$ is standard Gaussian white noise.

*Single Differentiation:* In the case of RCIE, let $n_{\rm c} = 1$, $n_{\rm f} = 6$, $R_\theta = 10^{-6} I_3$, $R_d = 10^{-5}$, $R_z = 1$, $\widetilde{V} = 10^{-2}$. In the case of HGO, the parameters values are the same as they are for single differentiation in Example 4.1. For a signal-to-noise ratio (SNR) of 40 dB ($D_2 = 0.00699945$), Figure 3 compares the signals estimated by SD/RCIE, SD/HGO, and SD/BDD with the true first derivative.

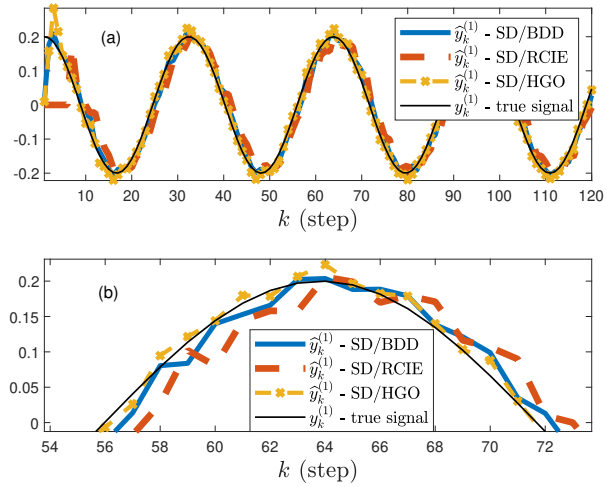


Fig. 3. *Example 4.2 Single Differentiation.* (a) The signals estimated by SD/RCIE, SD/HGO, and SD/BDD follow the true first derivative $y^{(1)}$ after an initial transient period. SD/HGO exhibits a longer transient period as compared to SD/RCIE. (b) A zoomed view of plot (a). At steady state, SD/HGO is more accurate than SD/RCIE and SD/BDD.

*Double Differentiation:* In the case of RCIE, let $n_{\rm c} = 18$, $n_{\rm f} = 4$, $R_\theta = 10^{-1} I_{37}$, $R_d = 10^{-6}$, $R_z = 1$, $\widetilde{V} = 10^{-5}$. In the case of HGO, the parameters values are the same as they

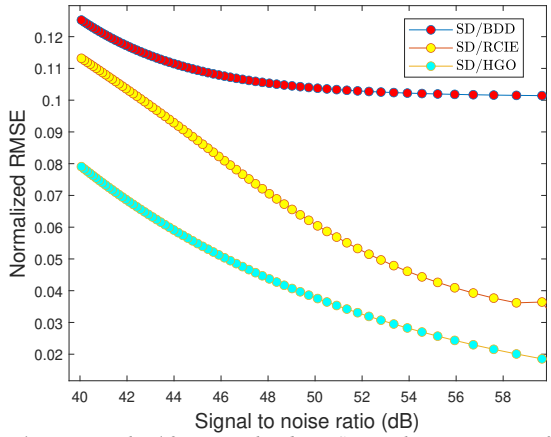


Fig. 4. *Example 4.2. Normalized RMSE in the estimation of the first derivative.* SD/HGO performs better than SD/RCIE and SD/BDD.

are for double differentiation in Example 4.1. For an SNR of 40 dB, Figure 5 compares the signals estimated by DD/RCIE, DD/HGO, and DD/BDD with the true second derivative.

In order to do quantitative comparison among the different methods, the normalized RMSE in the estimation of the single derivative and the double derivative is plotted in Figures 4 and 6, respectively, for SNRs in the range of 40 dB to 60 dB.

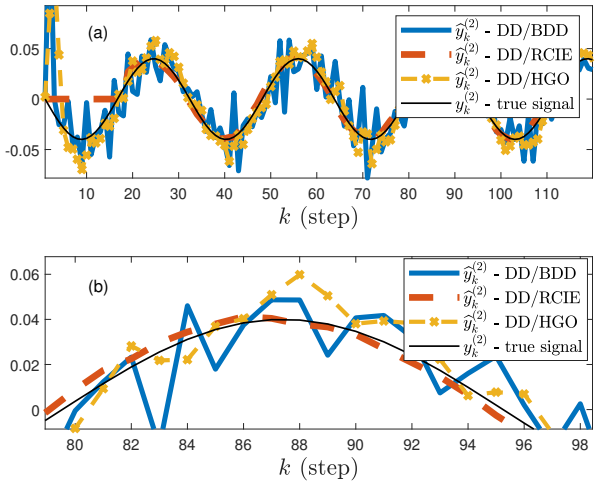


Fig. 5. *Example 4.2 Double Differentiation.* (a) The signal estimated by DD/RCIE follows the true second derivative $y^{(2)}$ after an initial transient period. Though the signals estimated by DD/HGO and DD/BDD follow the general trend of $y^{(2)}$, they are noisy. (b) A zoomed view of plot (a). At steady state, DD/RCIE is more accurate than DD/HGO and DD/BDD.

## V. DIFFERENTIATION OF EXPERIMENTAL DATA

RCIE, HGO, and BDD are applied to experimental position data of a small rover for estimating its velocity and acceleration. An OptiTrack camera sensor is used to collect the position data of the rover at a sample rate of 50 Hz. Figure 7 depicts the trajectory of the rover on the $x$-$y$ plane and the position data along the $x$-axis. Differentiation of the position data along the $x$-axis is done to obtain the velocity and the acceleration along the $x$-axis. Since the true velocity and the true acceleration of the rover are not known, it is not possible to evaluate the accuracy of the estimated signals.

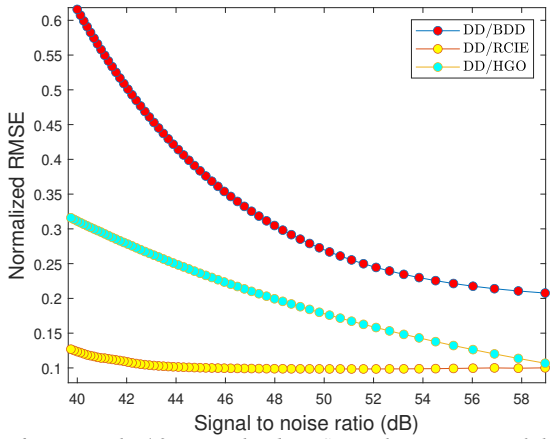


Fig. 6. *Example 4.2. Normalized RMSE in the estimation of the second derivative.* DD/RCIE performs better than DD/HGO and DD/BDD.

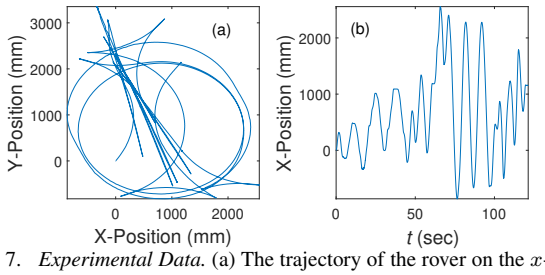


Fig. 7. *Experimental Data.* (a) The trajectory of the rover on the $x$-$y$ plane. (b) Position of the rover along $x$-axis versus time.

*Single Differentiation:* In the case of RCIE, let $n_c = 1$, $n_f = 6$, $R_\theta = 10^{-3} I_3, R_d = 10^{-5}, R_z = 1$, $\widetilde{V} = 10^{-4}$. In the case of HGO, let $n = 3$ in (27), (28), and (29), let $\alpha_1 = 3, \alpha_2 = 3$, $\alpha_3 = 1$. The parameter $\varepsilon$ is given an optimum value that renders the estimated signal smooth and follow the general trend of the signals estimated by RCIE and BDD. Figure 8 compares the signals estimated by SD/RCIE, SD/HGO, and SD/BDD.

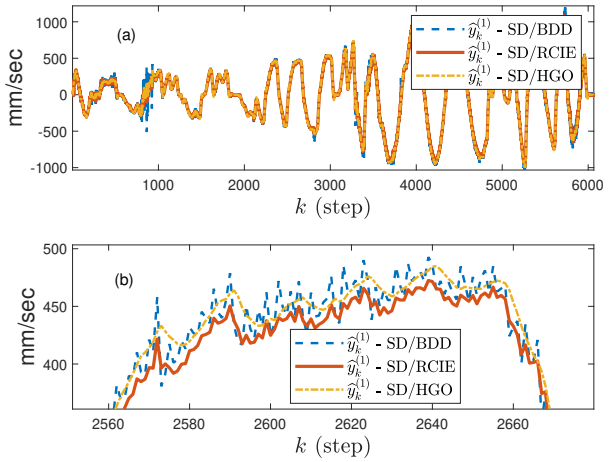


Fig. 8. *Single Differentiation of Experimental Data.* (a) shows the signals estimated by SD/RCIE, SD/HGO, and SD/BDD. (b) A zoomed view of plot (a). The signal estimated by SD/BDD is noisy, whereas the signals estimated by SD/RCIE and SD/HGO are reasonably smooth.

*Double Differentiation:* In the case of RCIE, let $n_c = 18$, $n_f = 4$, $R_\theta = 10^{-1} I_{37}, R_d = 10^{-6}, R_z = 1$, $\widetilde{V} = 10^{-5}$. In the case of HGO, let $n = 4$ in (27), (28), and (29), let $\alpha_1 = 8, \alpha_2 = 24$, $\alpha_3 = 32$, $\alpha_4 = 16$. The parameter $\varepsilon$ is chosen

in the same way as chosen for single differentiation. Figure 9 compares the signals estimated by DD/RCIE, DD/HGO, and DD/BDD.

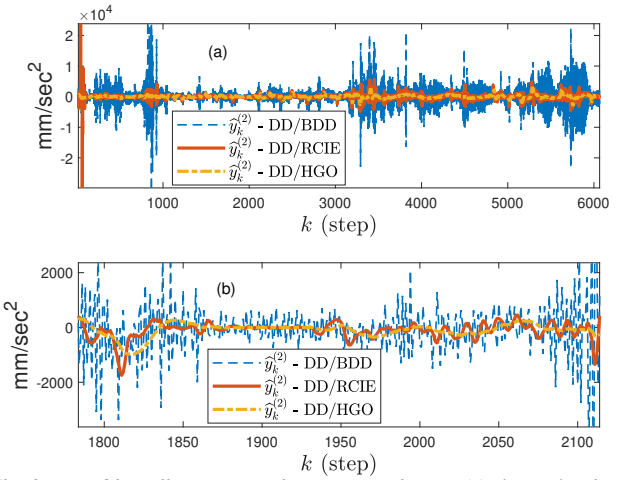


Fig. 9. *Double Differentiation of Experimental Data.* (a) shows the signals estimated by DD/RCIE, DD/HGO, and DD/BDD. (b) A zoomed view of plot (a). The signal estimated by DD/BDD is noisy, whereas the signals estimated by DD/RCIE and DD/HGO are reasonably smooth.

## VI. CONCLUSIONS

Numerical differentiation based on sampled data is a longstanding and difficult problem due to the effect of noise and the lack of a truth sensor. Using simulated data with simulated noise, retrospective cost input estimation (RCIE) was compared with high-gain observers (HGO), using backward difference differentation (BDD) as a baseline. Since the true signal is known, the accuracy of these methods could be compared, and it was shown that HGO outperforms RCIE for single differentiation, whereas RCIE outperforms HGO for double differentiation.

These methods were then applied to experimental data for which the true derivatives were unavailable. Consequently, it is not possible to definitively compare the accuracy of the different methods. Nevertheless, the estimated first and second derivatives suggest that the BDD estimates are noisy, as expected. However, both RCIE and HGO provide smooth estimates of the derivatives, which are physically reasonable. Future research will focus on the development of metrics that can be used to determine the relative accuracy of the RCIE and HGO estimates.

## ACKNOWLEDGMENTS

This research was supported by NSF grant CMMI 2031333 and Ford. The authors are grateful to Ruiyang Wang for providing the experimental data.

## REFERENCES

[1] R. Vilanova and A. Visioli, *PID Control in the Third Millennium: Lessons Learned and New Approaches*. Springer, 2012.
[2] K. Astrom and T. Hagglund, *Advanced PID Control*. ISA, 2006.
[3] J. Cullum, "Numerical Differentiation and Regularization," *SIAM J. Numer. Analysis*, vol. 8, pp. 254–265, 1971.
[4] R. Chartrand, "Numerical Differentiation of Noisy, Nonsmooth Data," *ISRN Appl. Math.*, 2011. Article ID 164564.
[5] I. Knowles and R. J. Renka, "Methods for Numerical Differentiation of Noisy Data," *Electron. J. Differ. Equ*, vol. 21, pp. 235–246, 2014.
[6] S. Diop, J. W. Grizzle, P. E. Moraal, and A. Stefanopoulou, "Interpolation and Numerical Differentiation for Observer Design," in *Proc. Amer. Contr. Conf.*, pp. 1329–1333, 1994.
[7] S. Diop, J. Grizzle, and F. Chaplais, "On Numerical Differentiation Algorithms for Nonlinear Estimation," in *Proc. IEEE Conf. Dec. Contr. (Cat. No.00CH37187)*, vol. 2, pp. 1133–1138, 2000.
[8] J. A. Walker, "Estimating Velocities and Accelerations of Animal Locomotion: A Simulation Experiment Comparing Numerical Differentiation Algorithms," *J. Exp. Biol.*, vol. 201, no. 7, pp. 981–995, 1998.
[9] A. M. Dabroom and H. K. Khalil, "Discrete-time Implementation of High-Gain Observers for Numerical Differentiation," *Int. J. Contr.*, vol. 72, no. 17, pp. 1523–1537, 1999.
[10] M. Mboup, C. Join, and M. Fliess, "A Revised Look at Numerical Differentiation with an Application to Nonlinear Feedback Control," in *Medi. Conf. Contr. Autom.*, pp. 1–6, 2007.
[11] M. Mboup, C. Join, and M. Fliess, "Numerical Differentiation with Annihilators in Noisy Environment," *Numer. Algor.*, vol. 50, pp. 439–467, 2009.
[12] M. Reichhartinger and S. Spurgeon, "An Arbitrary-Order Differentiator Design Paradigm with Adaptive Gains," *Int. J. Contr.*, vol. 91, no. 9, pp. 2028–2042, 2018.
[13] F. López-Caamal and J. A. Moreno, "Generalised Multivariable Supertwisting Algorithm," *Int. J. Robust Nonlinear Contr.*, vol. 29, no. 3, pp. 634–660, 2019.
[14] F. Van Breugel, J. N. Kutz, and B. W. Brunton, "Numerical Differentiation of Noisy Data: A Unifying Multi-Objective Optimization Framework," *IEEE Access*, vol. 8, pp. 196865–196877, 2020.
[15] M. Hou and R. J. Patton, "Input Observability and Input Reconstruction," *Automatica*, vol. 34, no. 6, pp. 789–794, 1998.
[16] M. Corless and J. Tu, "State and Input Estimation for a Class of Uncertain Systems," *Automatica*, vol. 34, no. 6, pp. 757–764, 1998.
[17] T. Floquet and J. P. Barbot, "State and Unknown Input Estimation for Linear Discrete-Time Systems," *Automatica*, vol. 42, pp. 1883–1889, 2006.
[18] S. Gillijns and B. De Moor, "Unbiased Minimum-Variance Input and State Estimation for Linear Discrete-Time Systems," *Automatica*, vol. 43, no. 1, pp. 111–116, 2007.
[19] R. Orjuela, B. Marx, J. Ragot, and D. Maquin, "On the Simultaneous State and Unknown Input Estimation of Complex Systems via a Multiple Model Strategy," *IET Contr. Theory Appl.*, vol. 3, no. 7, pp. 877–890, 2009.
[20] H. Palanthandalam-Madapusi and D. S. Bernstein, "A Subspace Algorithm for Simultaneous Identification and Input Reconstruction," *Int. J. Adapt. Contr. Sig. Proc.*, vol. 23, pp. 1053–1069, 2009.
[21] S. Kirtikar, H. Palanthandalam-Madapusi, E. Zattoni, and D. S. Bernstein, "$l$-Delay Input and Initial-State Reconstruction for Discrete-Time Linear Systems," *Circ. Syst. Sig. Proc.*, vol. 30, no. 1, pp. 233–262, 2011.
[22] H. Fang, R. A. de Callafon, and J. Cortes, "Simultaneous Input and State Estimation for Nonlinear Systems with Applications to Flow Field Estimation," *Automatica*, vol. 49, no. 9, pp. 2805–2812, 2013.
[23] S. Z. Yong, M. Zhu, and E. Frazzoli, "A Unified Filter for Simultaneous Input and State Estimation of Linear Discrete-Time Stochastic Systems," *Automatica*, vol. 63, pp. 321–329, 2016.
[24] P. Lu, E.-J. van Kampen, C. C. de Visser, and Q. Chu, "Framework for State and Unknown Input Estimation of Linear Time-Varying Systems," *Automatica*, vol. 73, pp. 145–154, 2016.
[25] C.-S. Hsieh, "Unbiased Minimum-Variance Input and State Estimation for Systems with Unknown Inputs: A System Reformation Approach," *Automatica*, vol. 84, pp. 236–240, 2017.
[26] S. Sanjeevini and D. S. Bernstein, "Minimal-Delay FIR Delayed Left Inverses for Systems with Zero Nonzero Zeros," *Sys. Contr. Lett.*, vol. 133, p. 104552, 2019.
[27] E. Naderi and K. Khorasani, "Unbiased Inversion-Based Fault Estimation of Systems with Non-Minimum Phase Fault-to-Output Dynamics," *IET Contr. Th. & Appl.*, vol. 13, no. 11, pp. 1629–1638, 2019.
[28] B. Alenezi, M. Zhang, S. Hui, and S. H. Zak, "Simultaneous Estimation of the State, Unknown Input, and Output Disturbance in Discrete-Time Linear Systems," *IEEE Trans. Autom. Contr.*, 2021.
[29] A. Ansari and D. S. Bernstein, "Input Estimation for Nonminimum-Phase Systems With Application to Acceleration Estimation for a Maneuvering Vehicle," *IEEE Trans. Contr. Syst. Techn.*, vol. 27, no. 4, pp. 1596–1607, 2019.