Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Analysis of the parameter estimate error when algebraic differentiators are used in the presence of disturbances ⋆

A. Othmane [*,**] **J. Rudolph** [**] **H. Mounier** [*]

[*] *Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systèmes, 91190, Gif-sur-Yvette, France (e-mail: {amine.othmane, hugues.mounier}@l2s.centralesupelec.fr)*
[**] *Chair of Systems Theory and Control Engineering Saarland University, Saarbrücken, Germany (e-mail: {a.othmane, j.rudolph}@lsr.uni-saarland.de)*

Abstract: The use of algebraic differentiators in the context of asymptotic continuous-time parameter estimation is discussed. The estimation problem is analyzed within a least squares optimization context. Bounds for the error stemming from high frequency disturbances and the approximation of the derivatives are derived. It is shown that with higher frequencies the error stemming from the disturbances decreases and that the filter parameters can be used to adjust the convergence of this error to zero. An observer with assignable error dynamics for the online estimation is also proposed. A simulation is carried out to evaluate the results and compare the proposed observer with the recursive solution of the least squares problem.

Keywords: Algebraic differentiators, parameter estimation, least squares algorithm, time-varying linear observers

1. INTRODUCTION

The problem of identifying parameters of a dynamic process from its operating data is an important task and is still receiving considerable attention. Among the various techniques that where developed, regression models combined with a least squares estimator have attracted much attention. In Durbin (1960) for example, this approach is developed for discrete time dynamical systems. However, in the presence of disturbed measurements, the least squares analysis may lead to inconsistent estimates as discussed in detail for example in Wald (1940). Different approaches have been proposed in the literature in order to overcome this problem. The instrumental variable approach for example, first developed in Reiersøl (1941); Durbin (1954); Young (1970), can be seen as a slight variation of the least squares solution of the linear regression overcoming the problem. However, as noted in Young (1979), the major problem with this approach is the generation of the instrumental variables themselves.

For the continuous-time estimation of process parameters using the least squares approach, the knowledge of the time derivatives of the measured signals is required and in the presence of disturbances represents an additional challenge. In Young (1979), a state variable filter is proposed which simultaneously filters the signals and provides filtered time derivatives which replace the exact but unknown derivatives.

In Mboup et al. (2007); Mboup et al. (2009), a family of numerical differentiators was introduced. The time derivative of a measured signal can be approximated using these estimators, which are time-invariant filters and operate on the non-differentiated signal. Their stability is implicitly guaranteed from their finite impulse response. In Liu (2011), it was shown that the algebraic differentiators outperform recent numerical algorithms based on higher order sliding modes and high-gain observers. For a systematic configuration of the five filter parameters the methods presented in Kiltz and Rudolph (2013); Kiltz (2017) can be applied such that desired properties of the amplitude spectrum of the derivative estimator are achieved. Two methods were presented: The first one allows the elimination of harmonic disturbances. In the second one, the derivative estimators are approximated as low-pass filters of arbitrary large order. Then, the parameters can be calculated from the desired cutoff frequency and the desired filter order.

The contribution of this article is to analyze the parameter estimate error when algebraic differentiators are used in the context of continuous-time parameter estimation using linear regression with a least squares approach. Bounds for the error stemming from the disturbances and the derivative approximation are derived using recent results from Kiltz (2017). High frequency measurement disturbances are considered. It is shown that the rate of convergence to zero of the bias stemming from these disturbances can be tuned using the parameters of the algebraic differentiators. Furthermore, an observer with assignable error dynamics and delayed measurements for the online estimation of the parameters using the approximated derivatives is

⋆ This work is supported by the "ADI 2018" project funded by the IDEX Paris-Saclay, ANR-11-IDEX-0003-0

presented and compared with the recursive solution of the least squares optimization problem in a simulation.

This article is structured as follows. In section 2, the notation used in the work is summarized. The algebraic differentiators introduced in Mboup et al. (2007); Mboup et al. (2009) are recalled in section 3. The results from Kiltz and Rudolph (2013); Kiltz (2017) are used to analyze their behavior in the presence of high frequency disturbances and error bounds are derived. In section 4, the use of these filters in the context of parameter estimation is discussed. First, the use of a least squares optimization problem is introduced and analyzed. Then, an observer with assignable error dynamics is proposed. In section 5, a simulation is performed and the effectiveness of the proposed approaches is shown.

2. NOTATION

For a real-valued function f defined on $\mathbb{R}$, the supremum norm of f is defined as $\|f\|_\infty = \sup_{t\in\mathbb{R}} |f(t)|$. For $\mathcal{S} \subset \mathbb{R}$ it holds that $\|f\|_{\infty,\mathcal{S}} = \sup_{t\in\mathcal{S}} |f(t)|$. The convolution of two real-valued functions f and g is written $f * g$. A function f is said to be of class $\mathcal{C}^n$, for n a non-negative integer, if the derivatives of f up to the order n exist and are continuous. In order to specify a $\mathcal{C}^n$ function on a domain $I \subseteq \mathbb{R}$, the notation $\mathcal{C}^n(I)$ is adopted. For a $\mathcal{C}^n$ function f, with n a non-negative integer, $f^{(i)}$, for $i = 0, \ldots, n$, denotes its i-th derivative and $f^{(0)}$ is the function f itself. Alternatively, Newtons notation for derivatives is also adopted for the first and the second derivative. The gamma function is denoted as Γ. The orthogonal Jacobi polynomial of degree i associated with the weight function

$$w^{(\alpha,\beta)}(\tau) = \begin{cases} (1-\tau)^\alpha(1+\tau)^\beta, & \tau \in [-1,1] \\ 0, & \text{else,} \end{cases}$$

with the real parameters $\alpha, \beta > -1$, is denoted $P_i^{(\alpha,\beta)}$. The definitions for class $\mathcal{K}$ and $\mathcal{KL}$ functions are borrowed from Khalil (2002).

3. ALGEBRAIC DIFFERENTIATORS

3.1 Time domain interpretation of algebraic differentiators

The algebraic derivative estimation methods introduced in Mboup et al. (2007); Mboup et al. (2009) were initially derived using differential algebraic manipulations of truncated Taylor series. Later works (Liu (2011); Kiltz (2017)) derived these filters using an approximation theoretical approach that yields a straightforward analysis of the filter characteristics, especially the estimation delay. Using this approach, the estimate of the n-th order derivative of a signal $t \mapsto y(t)$ denoted $\hat{y}^{(n)}$ can be approximated as

$$\hat{y}^{(n)}(t) = \int_0^T g_T^{(n)}(\tau) y(t-\tau)\,\mathrm{d}\tau, \tag{1}$$

with the filter kernel

$$g_T(t) = \frac{2}{T}\sum_{i=0}^{N} \frac{P_i^{(\alpha,\beta)}(\vartheta)}{\left\|P_i^{(\alpha,\beta)}\right\|^2} w^{(\alpha,\beta)}\left(\nu(t)\right) P_i^{(\alpha,\beta)}\left(\nu(t)\right).$$

In the latter equation

- $\nu(t) = 1 - \frac{2}{T}t$,
- $\|x\| = \sqrt{\langle x, x\rangle}$ is the norm induced by the inner product
$$\langle x, y\rangle = \int_{-1}^{1} w^{(\alpha,\beta)}(\tau)x(\tau)y(\tau)\,\mathrm{d}\tau,$$
- N is the degree of the polynomial approximating the signal $y^{(n)}$ in the time window $[t-T, t]$,
- T is the filter window length,
- and ϑ parameterizes the estimation delay as described in (3).

This approach yields a straightforward analysis of the estimation delay δ_t and the degree of exactness[1] γ which are given as (Kiltz (2017))

$$\gamma = \begin{cases} n+N+1, \text{ if } N = 0 \vee \vartheta = p_{N+1,k}^{(\alpha,\beta)} \\ n+N, \qquad \text{otherwise,} \end{cases} \tag{2}$$

$$\delta_t = \begin{cases} \frac{\alpha+1}{\alpha+\beta+2}T, & \text{if } N = 0 \\ \frac{1-\vartheta}{2}T, & \text{otherwise,} \end{cases} \tag{3}$$

with $p_{N+1,k}^{(\alpha,\beta)}$ the k-th zero of the Jacobi polynomial of degree $N+1$. In the sequel, h_T is a function defined as

$$h_T(t) = \begin{cases} g_T(t), & \text{if } t \in [0,T], \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

3.2 Error analysis

Let $\bar{y}$ be a disturbed measurement of y such that

$$\bar{y}(t) = y(t) + \eta(t), \tag{5}$$

where η represents the additive disturbance. The estimated derivative $\hat{y}^{(n)}$ is corrupted by three sources of errors (Mboup et al. (2007); Mboup et al. (2009); Kiltz (2017)):

- The error e_d stemming from the delay δ_t and defined as $e_\mathrm{d}(t) \coloneqq y^{(n)}(t-\delta_t) - y^{(n)}(t)$.
- The polynomial approximation yields a smoothing of fast changes in the signal. This error is denoted e_a and is defined as $e_\mathrm{a}(t) \coloneqq (h_T * y)(t) - y^{(n)}(t-\delta_t)$.
- The error due to the disturbance corrupting the measurements and denoted e_n. It is defined as
$$e_\mathrm{n} \coloneqq h_T^{(n)} * y - h_T^{(n)} * \bar{y} = h_T^{(n)} * \eta.$$

The errors e_a and e_n are of interest in this work and are studied in the following propositions.

Proposition 1. (Liu (2011); Kiltz (2017)).
Let $y \in \mathcal{C}^{n+m+1}(I)$, with $m \in \mathbb{N}$ and $I \subseteq \mathbb{R}$, γ and h_T the degree of exactness defined in (2) and the filter given in (4), respectively. Assume there exists $M_n \in \mathbb{R}_+^*$ such that $\left\|y^{(n+p+1)}\right\|_{\infty,I} \leqslant M_n$ with $p = \min\{m, \gamma - n\}$. Then, there exists a positive scalar M such that the approximation error satisfies

$$\|e_\mathrm{a}\|_{\infty,I} \leqslant \frac{M}{p!}\left(\frac{T}{2}\right)^{p+1} M_n\,,$$

for $T \to 0$.

Proposition 2. Let $\bar{y}$ given in (5) be a disturbed measurement of $y \in \mathcal{C}^{n+1+p}(I)$, $p, n \in \mathbb{N}$, $I \subseteq \mathbb{R}$ and h_T the

[1] The degree of exactness was introduced in Kiltz (2017) as the polynomial degree up to which the derivative estimation is exact. If $\gamma = 2$ for example, the first and second time derivatives of a polynomial signal of degree two are exact up to an estimation delay.

filter kernel defined in (4). The disturbance is assumed to be integrable and bounded by σ, i.e., $\sigma = \sup_{t\in I} \eta(t)$. Then, for all times $t \in I$ the error e_n stemming from the disturbance is bounded such that

$$\|e_\mathrm{n}\|_{\infty,I} \leqslant \frac{\sigma}{T^{n+1}} Q,$$

with $Q = T^{n+1} \int_0^T \left|g^{(n)}(\tau)\right| \mathrm{d}\tau < \infty$ independent of T.

Proof. The error stemming from the disturbance is given as

$$\begin{aligned} e_\mathrm{n}(t) &= (h_T^{(n)} * y)(t) - (h_T^{(n)} * \bar{y})(t) \\ &= (h_T^{(n)} * \eta)(t) \\ &= \int_0^T g_T^{(n)}(t-\tau)\eta(\tau)\,\mathrm{d}\tau. \end{aligned}$$

Note that

$$g^{(n)}(t) = \frac{2^{2n+1}}{T^{n+1}} w^{(\alpha-n,\beta-n)}(\nu(t)) \cdot \left(\sum_{i=0}^{N} \frac{(i+n)!\, P_i^{(\alpha,\beta)}(\vartheta)}{i! \left\| P_i^{(\alpha,\beta)} \right\|^2} P_{i+n}^{(\alpha-n,\beta-n)}(\nu(t)) \right),$$

with $\nu(t) = 1 - \frac{2}{T}t$. Let $Q = T^{n+1} \int_0^T \left|g^{(n)}(\tau)\right| \mathrm{d}\tau < \infty$ which is independent of T. Thus, the error can be bounded by

$$\begin{aligned} |e_\mathrm{n}(t)| &= \left| \int_0^T g_T^{(n)}(t-\tau)\eta(\tau)\,\mathrm{d}\tau \right| \\ &\leqslant \sigma \int_0^T \left| g_T^{(n)}(\tau) \right| \mathrm{d}\tau = \frac{\sigma}{T^{n+1}} Q, \end{aligned}$$

for all times t. Hence, the error satisfies $\|e_\mathrm{n}\|_{\infty,I} \leqslant \frac{\sigma}{T^{n+1}} Q$.

The error bounds for more general stochastic processes are derived in Liu (2011). However, it is stressed that these are only conservative bounds and do not take into account the low pass effect of these filters analyzed in detail in Kiltz and Rudolph (2013); Kiltz (2017).

The estimation error in the presence of high frequency disturbances is now analyzed using the results from Kiltz and Rudolph (2013); Kiltz (2017). For this purpose, a disturbance model introduced in Astolfi et al. (2016) for the analysis of the sensitivity of High-Gain observers to high frequency disturbances is used. The model is given as

$$\epsilon \dot{w}(t) = S w(t), \quad \eta(t) = P w(t), \tag{6}$$

where the matrices S and P take the form

$$S = \mathrm{blockdiag}(S_1, \ldots, S_m), \quad S_i = \begin{bmatrix} 0 & \omega_i \\ -\omega_i & 0 \end{bmatrix},$$
$$P = ((0\ p_1), (0\ p_2), \ldots, (0\ p_m)),$$

with $\omega_i \in \mathbb{R}_{>0}$, $p_i \in \mathbb{R}$, $i = 1, \ldots, m$, $m \in \mathbb{N}$, and the parameter $\epsilon \in (0,1)$ will be taken small in the forthcoming analysis. The initial conditions are denoted $w_0(t_0) = w_0$. System (6) can be conveniently seen as a generator of $m > 0$ harmonics at frequencies $\frac{w_i}{\epsilon} > 0$, $i = 1, \ldots, m$. The disturbance is then given as $\eta(t) = \sum_{i=1}^m p_i \cos(\frac{\omega_i}{\epsilon} t + \phi_i)$ where $\phi_i = \arccos(w_{0,i,2})$. The next proposition proposes a bound for the error stemming from a high frequency disturbance generated by the model given in (6).

Proposition 3. Let $\bar{y}$ given in (5) be a disturbed measurement of $y \in \mathcal{C}^n(I)$, $n \in \mathbb{N}$, $I \subseteq \mathbb{R}$, and h_T the filter kernel defined in (4). Assume the disturbance to be generated by the model in (6). Introduce $\omega_> = \max_{i\in\{1,\ldots,m\}} \omega_i$ and $\omega_< = \min_{i\in\{1,\ldots,m\}} \omega_i$. Then, for $\epsilon \to 0$ and $\min\{\alpha, \beta\} > n$ the following holds for the error stemming from the disturbance:

- If $N = 0$ and $\alpha = \beta$, there exists a function $\kappa_1 \in \mathcal{KL}$ such that
$$0 \leqslant \|e_\mathrm{n}\|_{\infty,I} \leqslant \kappa_1\left(\frac{1}{T}, \frac{\omega_>}{\epsilon}\right),$$
- otherwise there exist two functions $\kappa_2, \kappa_3 \in \mathcal{KL}$ such that
$$\kappa_2\left(\frac{1}{T}, \frac{\omega_<}{\epsilon}\right) \leqslant \|e_\mathrm{n}\|_{\infty,I} \leqslant \kappa_3\left(\frac{1}{T}, \frac{\omega_>}{\epsilon}\right).$$

Proof. As observed in earlier references (see for instance Mboup et al. (2009); Kiltz and Rudolph (2013); Kiltz (2017)), the algebraic differentiator in (1) is a linear time-invariant filter. Thus, the response of the filter to a sinusoidal signal with amplitude u and angular frequency ω_0 is a sinusoidal signal with an identical frequency and amplitude $U = u\omega_0^n \mathcal{G}_T(\omega_0)$, where $\mathcal{G}_T$ denotes the Fourier transform of g_T. It is shown in (Kiltz, 2017, chapter 3) that for $|\omega| \to \infty$ the function $\mathcal{G}_T$ satisfies

$$\frac{1}{(\omega T)^\mu} \mathcal{G}^- \leqslant |\mathcal{G}_T(\omega)| \leqslant \frac{1}{(\omega T)^\mu} \mathcal{G}^+,$$

with

$$\mathcal{G}^+ = \left| r^{(\alpha,\beta)} \right| + \left| s^{(\alpha,\beta)} \right|$$
$$\mathcal{G}^- = \left| \left| r^{(\alpha,\beta)} \right| - \left| s^{(\alpha,\beta)} \right| \right|$$

where

$$\begin{aligned} \mu &= \min\{\alpha+1, \beta+1\}, \\ r^{(\alpha,\beta)} &= \sum_{i=0}^{N} \frac{c_i^{(\alpha,\beta)}}{\Gamma(\mu+\kappa+i)} P_i^{(\mu-1,\mu+\kappa-1)}(\chi\vartheta), \\ s^{(\alpha,\beta)} &= \sum_{i=0}^{N} \frac{(-1)^i c_i^{(\alpha,\beta)}}{\Gamma(\mu+i)} P_i^{(\mu-1,\mu+\kappa-1)}(\chi\vartheta), \\ c_i^{(\mu,\kappa)} &= (2\mu+\kappa+2i-1)\Gamma(2\mu+\kappa+i-1), \\ \chi &= \begin{cases} 1, & \beta \geq \alpha, \\ -1, & \alpha > \beta, \end{cases}, \quad \kappa = |\alpha - \beta|. \end{aligned}$$

The model (6) generates a sinusoidal disturbance signal $\eta(t) = \sum_{i=0}^m p_i \cos(\frac{\omega_i}{\epsilon} t + \phi_i)$. For notational simplicity, let $\omega = (\omega_1, \ldots, \omega_m)$. From the above considerations it follows that

$$\begin{aligned} |e_\mathrm{n}(t)| &\leqslant \left| \sum_{i=0}^{m} \left(\frac{w_i}{\epsilon}\right)^n p_i \mathcal{G}_T\left(\frac{\omega_i}{\epsilon}\right) \right| \\ &\leqslant \sum_{i=0}^{m} \left(\frac{w_i}{\epsilon}\right)^n |p_i| \left| \mathcal{G}_T\left(\frac{\omega_i}{\epsilon}\right) \right|, \end{aligned}$$

and, for $\epsilon \to 0$,

$$b_\mathrm{l}(T,\omega) \leqslant \|e_\mathrm{n}\|_{\infty,I} \leqslant b_\mathrm{u}(T,\omega),$$

where

$$b_\mathrm{l}(T,\omega) = T^{-\mu} \mathcal{G}^- \sum_{i=0}^{m} |p_i| \left(\frac{w_i}{\epsilon}\right)^{n-\mu},$$

$$b_\mathrm{u}(T,\omega) = T^{-\mu} \mathcal{G}^+ \sum_{i=0}^{m} |p_i| \left(\frac{w_i}{\epsilon}\right)^{n-\mu}.$$

Using $\omega_>$ and $\omega_<$ it follows that

$$b_\mathrm{l}(T,\omega) \geq mT^{-\mu}\left(\frac{\omega_<}{\epsilon}\right)^{n-\mu}\mathcal{G}^- p_< =: \kappa_2\left(\frac{1}{T},\frac{\omega_<}{\epsilon}\right)$$

$$b_\mathrm{u}(T,\omega) \leqslant mT^{-\mu}\left(\frac{\omega_>}{\epsilon}\right)^{n-\mu}\mathcal{G}^+ p_> =: \kappa_3\left(\frac{1}{T},\frac{\omega_>}{\epsilon}\right),$$

with $p_< = \min_{i\in\{1,\dots,m\}}|p_i|$ and $p_> = \max_{i\in\{1,\dots,m\}}|p_i|$. Thus,

$$\kappa_2\left(\frac{1}{T},\frac{\omega_<}{\epsilon}\right) \leqslant \|e_\mathrm{n}\|_{\infty,I} \leqslant \kappa_3\left(\frac{1}{T},\frac{\omega_>}{\epsilon}\right).$$

It is straightforward to see that for $\mu > n$ the functions $\kappa_2, \kappa_3 \in \mathcal{KL}$. For $N = 0$ and $\alpha = \beta$ it holds that

$$\kappa_2\left(\frac{1}{T},\frac{\omega_<}{\epsilon}\right) = 0$$

$$\kappa_3\left(\frac{1}{T},\frac{\omega_>}{\epsilon}\right) = \frac{4^\mu\Gamma(\mu+1/2)}{\sqrt{\pi}T^\mu\omega_>^\mu} =: \kappa_1\left(\frac{1}{T},\frac{\omega_>}{\epsilon}\right).$$

4. PARAMETER IDENTIFICATION USING ALGEBRAIC DIFFERENTIATORS

4.1 Least squares parameter estimation

Introductory example Consider the system

$$\dot{x}(t) = ax(t) + bu(t), \tag{7}$$

where $a, b \in \mathbb{R}$ are unknown parameters, u is a known function of time and $y(t) = x(t) + \eta(t)$ is a disturbance corrupted measurement of the state. Assume $\eta(t)$ is generated by (6). The estimation of the parameters a and b using a least squares approach and the algebraic differentiators presented above is now discussed.

Assume the filter in (1) is used to estimate $\dot{y}$ and denote the estimate as $\hat{\dot{y}}$. Taking into account the estimation delay δ_t, equation (7) can then be rewritten as

$$\hat{\dot{y}}(t) = ay(t-\delta_t) + bu(t-\delta_t).$$

For the purpose of convergence analysis of the algorithms the latter equation is rewritten as

$$\hat{\dot{x}}(t) = ax(t-\delta_t) + bu(t-\delta_t) + e(t),$$

with $e(t) = \hat{\dot{\eta}}(t) + \eta(t-\delta_t)$, where $\hat{\dot{\eta}}$ represents the output of the filter when applied to the disturbance signal. Note that defining $\theta = [a\ b]^T$, $H(t) = [x(t-\delta_t)\ u(t-\delta_t)]^T$, and $\phi(t) = \hat{\dot{x}}(t) - e(t)$ yields $\phi(t) = H^T(t)\theta$. Assume that for all $t, t_0 \geq 0$ satisfying $t_0 < t$ the function

$$\psi(t) = \int_{t_0}^t e^{\lambda(\tau-t)}H^T(\tau)H(\tau)\,\mathrm{d}\tau, \tag{8}$$

with $\lambda > 0$, is non-zero. Let $\hat{\theta}(t)$ be the solution of the minimization problem

$$\min_\theta \int_{t_0}^t e^{\lambda(\tau-t)}\left(H^T(\tau)\theta - \phi(\tau)\right)^T\left(H^T(\tau)\theta - \phi(\tau)\right)\mathrm{d}\tau.$$

Then,

$$\hat{\theta}(t) = \frac{1}{\psi(t)}\int_{t_0}^t H(\tau)\phi(\tau)\,\mathrm{d}\tau$$

$$= \underbrace{\frac{1}{\psi(t)}\int_{t_0}^t H(\tau)\hat{\dot{x}}(\tau)\,\mathrm{d}\tau}_{\theta^*(t)} - \underbrace{\frac{1}{\psi(t)}\int_{t_0}^t H(\tau)e(\tau)\,\mathrm{d}\tau}_{\tilde{\theta}(t)},$$

i.e., the measurement disturbance yields a bias $\tilde{\theta}$ in the estimation of the parameters. A method where this bias vanishes is presented in the sequel.

In contrast to the approach presented above, where the filter was only used for the estimation of the numerical derivative of the measurement, it is applied to all signals in the system to get

$$\hat{\dot{y}}(t) = a\hat{y}(t) + b\hat{u}(t) \tag{9}$$

and equivalently

$$\hat{\dot{x}}(t) = a\hat{x}(t) + b\hat{u}(t) + \hat{e}(t), \tag{10}$$

with $\hat{e}(t) = \hat{\dot{\eta}}(t) + \hat{\eta}(t)$. Using the same arguments as in the proof of Proposition 3 it can be shown that the error satisfies

$$\|e\|_\infty \leqslant \kappa_\epsilon\left(\frac{1}{T},\frac{\omega_>}{\epsilon}\right) = c_1T^{-\mu}\left|\left(\frac{\omega_>}{\epsilon}\right)^{1-\mu}\right|\left|1+\left(\frac{\omega_>}{\epsilon}\right)^{-1}\right|$$

with $\omega_> = \max_{i\in\{1,\dots,m\}}\omega_i$, $p_> = \max_{i\in\{1,\dots,m\}}p_i$, c_1 some constant depending on the filter parameters and the disturbance amplitude, and $\mu = \min\{\alpha+1, \beta+1\}$. Note that $\lim_{\epsilon\to 0}\kappa_\epsilon(\frac{\omega_>}{\epsilon}) = 0$ when $\mu > 1$. Estimating the parameters a and b from (10) using a least squares approach yields a bias $\tilde{\theta}$ satisfying

$$\begin{aligned}\left\|\tilde{\theta}\right\|_{\infty,I} &= \sup_{t\in I}\left|\frac{1}{\psi(t)}\int_{t_0}^t H(\tau)e(\tau)\,\mathrm{d}\tau\right| \\ &\leqslant \sup_{t\in I}\left|\frac{1}{\psi(t)}\right|\int_{t_0}^t |H(\tau)|\,|e(\tau)|\,\mathrm{d}\tau \\ &\leqslant \kappa_\epsilon\left(\frac{1}{T},\frac{\omega_>}{\epsilon}\right)\sup_{t\in I}\left|\frac{1}{\psi(t)}\right|\int_{t_0}^t|H(\tau)|\,\mathrm{d}\tau \\ &= \kappa_{\tilde{\theta}}\left(\frac{1}{T},\frac{\omega_>}{\epsilon}\right),\end{aligned}$$

for all $t \in I = [0\ t^*]$, with t^* bounded and strictly positive. If the input u is bounded, it can be concluded that $\lim_{\epsilon\to 0}\kappa_{\tilde{\theta}}\left(\frac{1}{T},\frac{\omega_>}{\epsilon}\right) = 0$, i.e., the identification bias stemming from the disturbances vanishes when $\epsilon \to 0$. The convergence speed of the error to zero can be adjusted using $\mu = \min\{\alpha+1, \beta+1\}$.

In the previous analysis only the effect of the disturbance on the estimation was taken into account. However, applying the algebraic differentiators also results in an error stemming from the polynomial approximation of the signals in the filter window. This is analyzed in the forthcoming paragraph.

For notational simplicity the disturbance is assumed to be identical to zero, i.e., $\eta(t) = 0$ for all $t \geq 0$. Define $e_x(t) = x(t-\delta_t) - \hat{x}(t)$ and $e_u(t) = u(t-\delta_t) - \hat{u}(t)$. Recalling the estimation equation (9)

$$\hat{\dot{x}}(t) = a\hat{x}(t) + b\hat{u}(t)$$

used earlier, yields

$$\dot{x}(t-\delta_t) = ax(t-\delta_t) + bu(t-\delta_t) + \underbrace{\dot{e}_x(t) - ae_x(t) - be_u(t)}_{e_g(t)}.$$

Thus, using the algebraic differentiators yields an estimation bias $e_g(t)$ compared to the use of the exact values. Assume that $x \in \mathcal{C}^{m_x+2}(I)$, $u \in \mathcal{C}^{m_u+1}(I)$, $I \subseteq \mathbb{R}$, and that these functions are bounded. Let $e_\theta(t) = \theta(t) - \hat{\theta}(t)$, $H(t) = [x(t)\ u(t)]^T$ and $\psi(t) = \int_{t_0}^t e^{\lambda(\tau-t)}H^T(\tau)H(\tau)\,\mathrm{d}\tau$

with $t_0 < t$. Assume $\psi(t) \neq 0$ for all times $t \in I$. The estimation error due to the use of the algebraic differentiators satisfies

$$\begin{aligned}\|e_\theta\|_{\infty,I} &= \sup_{t\in I}\left|\frac{1}{\phi(t)}\int_{t_0}^{t} H(\tau)e_g(\tau)\,\mathrm{d}\tau\right| \\ &\leqslant \sup_{t\in I}\left|\frac{1}{\phi(t)}\right|\int_{t_0}^{t}|H(\tau)|\,|e_g(\tau)|\,\mathrm{d}\tau \\ &\leqslant \|e_g\|_{\infty,I}\sup_{t\in I}\left|\frac{1}{\phi(t)}\right|\int_{t_0}^{t}|H(\tau)|\,\mathrm{d}\tau.\end{aligned}$$

From Proposition 1 it follows that

$$\|e_g\|_{\infty,I} = \mathcal{O}\left(\left(\frac{T}{2}\right)^{p+1}\right),$$

with $p = \min\{m_x, m_u, \gamma - 1\}$. Thus, the estimation error due to the use of the algebraic differentiators satisfies

$$\|e_\theta\|_{\infty,I} \leqslant \mathcal{O}\left(\left(\frac{T}{2}\right)^{p+1}\right).$$

Generalization The observation made with the simple first order system in (7) can be generalized for higher order systems as stated in the following proposition.

Proposition 4. Let n and m be positive integers, $u \in \mathcal{C}^m(I)$, $I \subseteq \mathbb{R}$, a known signal, h_T the filter defined in (4), x a $\mathcal{C}^n(I)$ function and $\bar{y} = x + \eta$ a disturbed measurement where the disturbance η is generated by the model in (6) with the parameter ϵ. Introduce $\hat{y}^{(i)} = h_T^{(i)} * \bar{y}$ and assume that u and x satisfy the linear relationship

$$x^{(n)}(t) + \sum_{i=0}^{n-1} a_i x^{(i)}(t) = \sum_{i=0}^{m} b_i u^{(i)}(t), \quad a_i,\, b_i \in \mathbb{R}. \quad (11)$$

Let λ be a positive scalar and define

$$\begin{aligned} H_x(t) &= \left[x(t), \ldots, x^{(n-1)}(t), u(t), \ldots, u^{(m)}(t)\right]^T, \\ H_{\bar{y}}(t) &= \left[\hat{y}(t), \ldots, \hat{y}^{(n-1)}(t), \hat{u}(t), \ldots, \hat{u}^{(m)}(t)\right]^T, \\ \theta &= [a_0, \ldots, a_{n-1}, b_0, \ldots, b_m]^T, \\ \psi_x(t) &= \int_{t_0}^{t} e^{\lambda(\tau - t)} H_x^T(\tau) H_x(\tau)\,\mathrm{d}\tau, \\ \psi_{\bar{y}}(t) &= \int_{t_0}^{t} e^{\lambda(\tau - t)} H_{\bar{y}}^T(\tau) H_{\bar{y}}(\tau)\,\mathrm{d}\tau. \end{aligned}$$

Assume the matrices $\psi_x(t)$ and $\psi_{\bar{y}}(t)$ are invertible for all $t \in I$. Denote by $\theta_x(t)$ and $\theta_{\bar{y}}(t)$ the solution of the following optimization problems

$$\min_\theta \int_{t_0}^{t} e^{\lambda(\tau - t)} \left(H_x^T(\tau)\theta - x^{(n)}(\tau)\right)^2 \mathrm{d}\tau,$$

$$\min_\theta \int_{t_0}^{t} e^{\lambda(\tau - t)} \left(H_{\bar{y}}^T(\tau)\theta - \hat{y}^{(n)}(\tau)\right)^2 \mathrm{d}\tau,$$

respectively. Assume further that $\lim_{t\to\infty} \theta_x(t) = \theta$. The estimated parameters using the measurements satisfy

$$\|\theta_x - \theta_{\bar{y}}\|_{\infty,I} \leqslant \kappa_g\left(\left(\frac{T}{2}\right)^{p+1}\right) + \kappa_\epsilon\left(\frac{1}{T}, \frac{\omega_>}{\epsilon}\right),$$

with $p = \min\{m, n, \gamma - 1\}$ and κ_g, κ_ϵ class $\mathcal{K}_\infty$ and class $\mathcal{KL}$ functions, respectively.

Proof. The proof follows from the considerations presented in the previous example.

4.2 Observer with assignable error dynamics

Consider the general model in (11) and the notation introduced in Proposition 4. First, for simplicity, assume that $x^{(i)}(t)$, $i = 0, \ldots, n$ and $u^{(i)}(t)$, $i = 0, \ldots, m$, are known for all times t. Introduce $\theta(t) = [a_0, \ldots, a_{n-1}, b_0, \ldots, b_m]^T$. The functions θ, H_x and $x^{(n)}$ satisfy

$$\begin{aligned} \dot{\theta}(t) &= 0 \\ x^{(n)}(t) &= H_x^T(t)\theta(t). \end{aligned}$$

Let

$$z(t) = \begin{bmatrix} x^{(n)}(t) \\ x^{(n)}(t-\delta_1) \\ \vdots \\ x^{(n)}(t-\delta_\kappa) \end{bmatrix} = \underbrace{\begin{bmatrix} H_x^T(t) \\ H_x^T(t-\delta_1) \\ \vdots \\ H_x^T(t-\delta_\kappa) \end{bmatrix}}_{=C(t)} \theta(t), \quad \kappa = n + m,$$

with the delays δ_i, $i \in \{1, \ldots, \kappa\}$ chosen such that $\det(C(t))$ is not identically zero for all times t. The parameters can be estimated using the observer

$$\dot{\hat{\theta}}(t) = L(t)\left(C(t)\hat{\theta}(t) - z(t)\right), \quad \hat{\theta}(0) \in \mathbb{R},$$

with $L(t)$ the solution of the equation

$$L(t)C(t) = \begin{cases} (\det(C(t)))^2 A, & \text{if } |\det(C(t))| < d \\ \rho A, & \text{otherwise,} \end{cases} \quad (12)$$

where A is an arbitrary matrix whose eigenvalues lie in the left half plane, and d and ρ are two strictly positive parameters. The asymptotic convergence of the estimates to the true parameters can be shown using the Lyapunov function $V(t) = e_\theta^T(t)e_\theta(t)$, with $e_\theta(t) = \theta(t) - \hat{\theta}(t)$.

When only the measurement $y(t) = x(t) + \eta(t)$, with the additive disturbance $\eta(t)$ satisfying (6), is available, an algebraic differentiator can be used and the parameters can be estimated by

$$\dot{\hat{\theta}}(t) = \hat{L}(t)\left(\hat{C}(t)\hat{\theta} - \hat{z}(t)\right), \quad \hat{\theta}(0) \in \mathbb{R},$$

with

$$\hat{C}(t) = \begin{bmatrix} H_{\bar{y}}^T(t) \\ H_{\bar{y}}^T(t-\delta_1) \\ \vdots \\ H_{\bar{y}}^T(t-\delta_\kappa) \end{bmatrix}, \quad \hat{z}(t) = (h_T^{(n)} * z)(t),$$

and $\hat{L}(t)$ the solution of (12) when the function C is replaced by $\hat{C}$. The estimate error derived in the last section for the least squares approaches can be derived for this observer in a similar manner.

5. EXAMPLE: PARAMETER ESTIMATION OF A 4-TH ORDER LINEAR SYSTEM

Consider the 4-th order linear system defined as

$$\dot{\xi}(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix} \xi(t) + \begin{bmatrix} 0 \\ \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix} u(t)$$

with the output $y(t) = \xi_1(t)$. The numerical values of the parameters are summarized in Tab. 1.

Table 1. Numerical values of the parameters used in the simulation.

γ_1	γ_2	γ_3	ζ_1	ζ_2	ζ_3
-16	-15.4	-1.2	0.5	-0.6	-0.1

The input and measurement disturbance are chosen to be a superposition of harmonics such that

$$u(t) = \sum_{i=1}^{6} \frac{u_i}{i} \left(\sin\left(\frac{\omega}{i} t\right) + \sin(\pi t) \right),$$

$$\eta(t) = \sum_{i=0}^{5} a_i \sin(\omega_i t + \phi_i),$$

with $u_i = 3$, $\omega = 10$, $a_i \in (0, 0.005)$ and $\omega_i \in [500, 5000]$ for $i \in \{0, \dots, 5\}$. The frequencies, amplitudes, and shifts of the disturbance and the input are chosen randomly using a uniform distribution. The signal-to-noise ratio is equal to 36 dB. In order to be able to compare the observer to the least squares estimator, the estimation of the parameters starts only for times greater than the highest delay used in the observer. The numerical values of the parameters of the algebraic differentiator are

$$N = 2, \quad T = 1.5, \quad \alpha = 5, \quad \beta = 7.5,$$

and ϑ is chosen to be the greatest zero of the Jacobi-polynomial $P_{N+1}^{(\alpha,\beta)}$. The parameters d and ρ of the observer are chosen to be equal to 10 and 1, respectively, and A is the diagonal matrix

$$A = -\mathrm{diag}([[2.4,\ 8,\ 3.2,\ 4,\ 0.8,\ 8]])$$

and the delays are

$$\delta_1 = 0.1,\ \delta_2 = 0.33,\ \delta_3 = 0.53,\ \delta_4 = 0.87,\ \delta_5 = 1.47.$$

The least squares optimization problem is solved recursively to show the applicability of the approach in the context of online parameter estimation. The parameter $\lambda = 0.98$. The observer is implemented using the backward Euler method. The sampling period is equal to 0.01.

The simulation results given in Fig. 1 show clearly that the estimated parameters converge to the true parameters. During the transient phase, the results of the observer show less oscillations and overshoots than those of the least squares estimator. Varying its parameter λ introduced in (8) does not significantly affect this observation.

6. CONCLUSION

The use of algebraic differentiators in the context of asymptotic parameter estimation has been discussed. The estimation error when a least squares approach is used has been analyzed using recent results from Kiltz (2017). More specifically, the bias stemming from high frequency disturbances has been considered. It has been shown that it decreases with higher disturbance frequencies. The parameters of the filters can be used to tune the convergence rate of the bias to zero. For the online estimation, an observer with assignable error dynamics has also been proposed and compared to the recursive solution of the least squares problem. The observer shows better results in the simulation. Further research will be conducted on the analysis of the error stemming from the discretization of the approaches presented here. A systematic methodology for the choice of the delays must also be discussed.

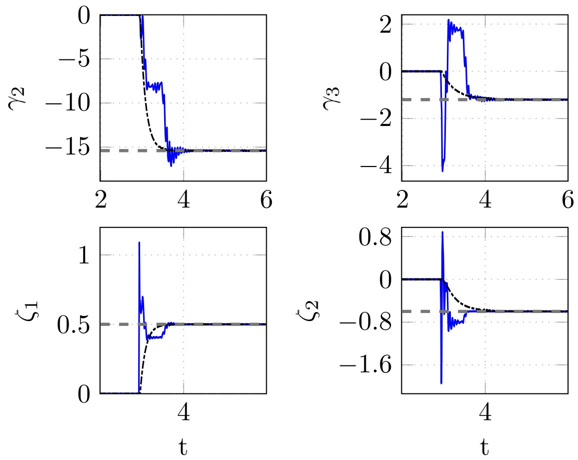

Fig. 1. The estimated parameters: ——— depicts the estimation results with the recursive least squares approach and - - - - when the proposed observer is used. The true parameters are given by - - - - - .

REFERENCES

Astolfi, D., Marconi, L., Praly, L., and Teel, A. (2016). Sensitivity to high-frequency measurement noise of nonlinear high-gain observers. *IFAC-PapersOnLine*, 49(18), 862 – 866.

Durbin, J. (1954). Errors in variables. *Revue de l'institut international de statistique / Review of the international statistical institute*, 22(1/3), 23–32.

Durbin, J. (1960). Estimation of parameters in time-series regression models. *Journal of the royal statistical society: Series B (Methodological)*, 22(1), 139–153.

Khalil, H. (2002). *Nonlinear systems*. Pearson Education. Prentice Hall.

Kiltz, L. (2017). *Algebraische Ableitungsschätzer in Theorie und Anwendung*. Ph.D. thesis, Saarland University.

Kiltz, L. and Rudolph, J. (2013). Parametrization of algebraic numerical differentiators to achieve desired filter characteristics. In *52nd IEEE Conference on Decision and Control*, 7010–7015. Florenz, Italy.

Liu, D.Y. (2011). *Analyse d'erreurs d'estimateurs des dérivées de signaux bruités et applications*. Ph.D. thesis, Université des Sciences et Technologies de Lille - Lille 1.

Mboup, M., Join, C., and Fliess, M. (2007). A revised look at numerical differentiation with an application to nonlinear feedback control. In *15th Mediterranean Conference on Control & Automation*, 1–6. Athens. Greece.

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

Reiersøl, O. (1941). Confluence analysis by means of lag moments and other methods of confluence analysis. *Econometrica*, 9(1), 1.

Wald, A. (1940). The fitting of straight lines if both variables are subject to error. *The Annals of Mathematical Statistics*, 11(3), 284–300.

Young, P. (1979). Parameter estimation for continuous-time models - a survey. *IFAC Proceedings Volumes*, 12(8), 17–41.

Young, P. (1970). An instrumental variable method for real-time identification of a noisy process. *Automatica*, 6(2), 271–287.