Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Directional Sensitivity of Least-squares State Estimators

Alexander Medvedev and Hannu Toivonen

Abstract— **Least-squares state estimators present an alternative to Luenberger observers and yield an exact (deadbeat) estimate of the state vector of a dynamic system as an optimal solution to a least-squares problem in some vector or functional space. Sensitivity of these estimators to structured uncertainty in the system matrix of the plant is studied in a common for continuous and discrete case framework using the Fréchet derivative. It is shown that the state estimation error caused by the plant model mismatch is proportional to the Fréchet derivative of the symbol of the parametrization operator used for the estimator implementation, evaluated for the nominal value of the system matrix. For the special case of state estimation in a single-tone continuous oscillator, the crucial impact of the parametrization operator choice on the observer sensitivity to plant model uncertainty is investigated in detail.**

I. Introduction

Least-squares state estimators present an alternative to Luenberger observers and yield an estimate of the state of a dynamic system as a solution to a least-squares problem in some suitable space. In this way, the very approach guarantees optimality of the produced estimate by minimizing a quadratic loss function. Vector spaces in both continuous [1] and discrete case [2] as well as functional (Banach) spaces [3] have been considered in deterministic and stochastic frameworks.

To make practical use of a least-squares estimator, the information about system's inputs and outputs should be collected at a finite time interval. Therefore, such observers are often termed as finite memory (or finite impulse response) observers or filters. Besides, in a noise-free case of an observable linear certain system, the least-squares estimate is exact, i.e. the estimation error converges to zero after a finite time has elapsed and thus exhibits deadbeat performance. In a stochastic framework, this property is equivalent to unbiasedness of the estimate, [4], [2].

Similar to their infinite memory counterparts, least-squares state estimators appear often e.g. in decoupling and disturbance rejection controllers [5], fault detection and isolation [6], control of time-delay systems [7]. In contrast to the Luenberger observer and the Kalman filter, the least-squares observers are used to limit estimate divergence due to process uncertainty by means of finite observer memory instead of robust feedback. This is also similar to the moving horizon approach to state estimation where limited memory is used to control the computational complexity of the underlying optimization problem, [8].

Since the pioneering work of Gilchrist [1], state estimators for continuous systems converging in finite time have been re-discovered many times. Most recent contributions to the field are [9] and [10], followed up by [11].

As mentioned above, all least-squares observers are stable due to their finite memory and optimal, by design, in the sense of a quadratic criterium. However, sensitivity of least-squares observers is seldom addressed. A special case of least-squares state estimation in a harmonic oscillator is analyzed in [12] demonstrating that observer robustness against oscillation frequency variation can be unacceptably low, especially when time delays are chosen for observer implementation.

In continuous time, the theory of least-squares state estimation based on the notion of pseudodifferential operator [13] provides necessary and sufficient existence conditions of the observers. It also stipulates that the observer properties are mainly defined by the pseudodifferential operator chosen for the observer parametrization.

Three particular operators have been treated so far. The differential operator is the classical one, see [14]. The time delay operator is though a more reasonable choice since it can be implemented in practice, see [15], [16]. The third operator, the sliding-window convolution operator, has been found to possess some beneficial properties when it comes to disturbance attenuation, [6], [7].

Corresponding analysis of the least-squares estimator in discrete time is not readily available. This can to some extent be explained by the fact that any discrete least-squares observer can be written as a finite impulse response filter. Thus, using different implementation operators along the lines of the continuous theory becomes a pure parametrization issue, see [17]. Nevertheless, parametrization is known to play a significant role in robustness properties of controllers and observers.

The fact that a least-squares observer is structurally stable does not necessarily guarantee good robustness properties of the state estimate. The lack of feedback in least-squares observers makes them vulnerable to parameter uncertainty. At the same time, the dynamic complexity of the least-squares observers can be easily increased, thanks to their filter-bank structure, which property is presumably helpful in shaping sensitivity functions.

In this paper, it is suggested that parameter sensitivity properties of the least-squares observers, both in discrete and continuous time, can suitably be analyzed via matrix function sensitivity theory, see [18], [19], [20], based on

The first author was partially supported by Swedish Research Council

A. Medvedev is with Information Technology, Uppsala University, P. O. Box 337, SE-751 05, SWEDEN, alexander.medvedev@it.uu.se

H. Toivonen is with Department of Information Technologies, Faculty of Technology, Åbo Akademi, FIN-20520 Åbo, Finland. hannu.toivonen@abo.fi

the Fréchet derivative. The paper is composed as follows. First the notion of continuous least-squares state estimator is revisited and followed up with a corresponding theory for the discrete time case. The sensitivity of the estimation error to structured perturbation in the plant system matrix are then studied in a common for discrete and continuous time formalism. Next, Fréchet derivatives for relevant parameterizations operators are evaluated to get insight into how free design parameters can be utilized for achieving higher observer robustness. Furthermore, the general theory is applied to the case of state estimation of a single-tone continuous oscillator, yielding recommendations on how parameters in two suitable operators have to be selected to achieve low state estimate sensitivity to oscillator frequency variations.

II. Least-squares state estimator

A. Continuous time

In this section, results necessary for further exposition on continuous least-squares observers originally presented in [13] are summarized. Only systems without exogenous signals are treated in the paper for simplicity of notation, although generalization to regular linear time-invariant systems is straightforward. Consider the autonomous system

$$\begin{aligned} \dot{x}(t) &= Ax(t) \\ y(t) &= Cx(t) \end{aligned} \tag{1}$$

where $x(t) \in \mathrm{R}^n$ is the state vector with the initial condition $x(0) = x_0$, $y(t) \in \mathrm{R}^\ell$ is the output vector, the eigenvalues of A comprise its spectrum $\sigma(A) = \{\mu_1, \ldots, \mu_n\}$.

Let the operator $(P\cdot)(\lambda; t)$, depending on parameter $\lambda \in \Lambda$ be defined via the Laplace inversion integral

$$(Pv)(\lambda; t) = \frac{1}{2\pi j} \int_{c-\infty}^{c+\infty} p(\lambda, s) V(s) e^{st}\, ds \tag{2}$$

where $V(s) = (\mathcal{L}v)(s)$ (i. e. the Laplace transform of $v(t)$), c is a suitable real constant and Λ is a nonempty real set of cardinality k. In the context of pseudodifferential operators, $p(\lambda, s)$ is called the symbol of the operator $(P\cdot)(\lambda; t)$.

Suppose that $p(\lambda, s)$ is analytic on $\sigma(A)$. Moreover, let the following two conditions hold.

Assumption 1 There exist real constants $c, \tau > 0$ such that the inequality

$$|p(\lambda, c + Re^{j\varphi})| \leq e^{-(c+R\cos\varphi)\tau}; \frac{\pi}{2} \leq \varphi \leq \frac{2\pi}{3} \tag{3}$$

holds for any large enough $R > 0$.

Condition (3) specifies the class of operators that can be used in least-squares state estimation by imposing an exponential decay of the operator symbol at infinity.

Assumption 2 For any given $\lambda \in \Lambda$ and $\mu_i \in \sigma(A)$, it holds that

(i) $p(\mu_i) \neq p(\mu_j)$ if $\mu_i \neq \mu_j$

(ii) $\frac{dp(s)}{ds} \mid_{s=\mu_i} \neq 0$ for every eigenvalue μ_i with height of the Jordan block greater than 1.

The above assumption is necessary to guarantee the preservation of all invariant subspaces of A under the transformation $p(\lambda, A)$.

Consider now the following observer

$$\hat{x}(t) = \mathcal{V}^{-1} \sum_{\lambda_i \in \Lambda} p(\lambda_i, A)^T C^T (Py)(\lambda_i; t) \tag{4}$$

where

$$\mathcal{V} = \sum_{\lambda_i \in \Lambda} p(\lambda_i, A)^T C^T C p(\lambda_i, A)$$

When Assumptions 1,2 are fulfilled, observer (4) possesses a deadbeat performance, in the sense that $e(t) = x(t) - \hat{x}(t) \equiv 0;\ t \geq \tau$ for any x_0. Furthermore, the minimal number of $\lambda_i \in \Lambda$ for which the observer always exists is equal to n, see [13] [6].

Naturally, deadbeat performance (or finite impulse response) in continuous time can only be achieved by means of an infinite-dimensional $p(\lambda, \cdot)$, which is of course a complication.

As demonstrated in [6], there is always a state-space realization of (1) such that the gramian matrix $\mathcal{V}$ is equal to a unit matrix. Thus, the least-squares observer can be, without loss of generality, written as

$$\hat{x}(t) = \sum_{\lambda_i \in \Lambda} p(\lambda_i, A)^T C^T (Py)(\lambda_i; t) \tag{5}$$

In the sequel, the system equations are assumed to be in a $\mathcal{V}$-balanced realization, if not stated otherwise.

B. Discrete time

In a discrete time framework, the issue of least-squares state estimation is much simpler because it does not involve infinite dimensional systems. Indeed, infinite dimensional dynamics have to be used in continuous time in order to obtain a finite memory in the observer. In discrete time, it suffices to consider operators that can be written as a finite polynomial of the discrete delay operator. The following result is a discrete time version of Lemma 1 in [13].

Lemma 1: Consider a discrete vector sequence $x(t) \in \mathrm{R}^n, t = 0, \ldots, \infty$

$$\begin{aligned} x(t) &= A^t x_0 : t = 0, \ldots, \infty \\ x_0 &= x(0) \end{aligned}$$

and a scalar transfer function $p(z)$ that is analytic on $\sigma(A)$.

Let $\tilde{x}(t) \in \mathrm{R}^n, t = 0, \ldots, \infty$ be a sequence whose Z-transform is

$$\tilde{X}(z) = p(z)X(z)$$

Then the following relationship holds

$$\tilde{x}(t) = p(A)x(t)$$

Proof: Omitted. ■

Notably, in discrete time, there is no restriction on how fast $|p(z)|$ decays at infinity, compared to the continuous time case. Otherwise, all the results obtained in [13] for the continuous least-squares observers are still valid in discrete time. They can be summarized for the purpose of this paper in the following manner.

Proposition 1: Consider an autonomous discrete system

$$\begin{aligned} x(t+1) &= Ax(t) \\ y(t) &= Cx(t) \end{aligned} \tag{6}$$

where $x \in \mathrm{R}^n$ is the state vector, $y \in \mathrm{R}^l$ is the output vector and (A, C) is an observable pair. Let a scalar transfer function $p(\lambda, z)$ be analytic on $\sigma(A)$ for all $\lambda \in \Lambda$. Suppose that for any given $\lambda \in \Lambda$ and $\mu_i \in \sigma(A)$ the conditions of Assumption 2 are fulfilled. Then the least-squares observer is given by (5). If, furthermore, for some constants a_i,$i = 0, \ldots, \tau$

$$p(\lambda, z) = \sum_{i=0}^{\tau} a_i z^{-i}$$

then (5) exhibits deadbeat performance in the sense that $e(t) = x(t) - \hat{x}(t) \equiv 0$; $t \geq \tau$ for any x_0.

Proof: Follows from Lemma 1 proceeding along the lines of Theorem 1 in [13]. ∎

A nice property of the least-squares state estimators is that the expression for the state estimate has the same form of equation (5), no matter whether continuous or discrete time framework is meant. This makes it possible to study the robustness properties of the estimators in both cases at the same time.

III. Sensitivity to system matrix uncertainty

A natural way of investigating the robustness properties of (5) is to take advantage of the Fréchet derivative that is widely applied in matrix function sensitivity theory. Indeed, a model uncertainty in the system matrix A gives rise to a corresponding perturbation of $p(\cdot, A)$ in (5). Achieving low sensitivity of $p(\cdot, A)$ with respect to a certain class of perturbations of A by choosing appropriate Λ yields local robustness of the least-squares state estimator against this particular class of perturbations.

Let a matrix function $p(A)$ be continuously differentiable at A in the sense that there exists a linear matrix operator $L_p(\cdot)$ such that for any matrix E

$$\lim_{\delta \to 0} \frac{p(A + \delta E) - p(A)}{\delta} = L_p(A, E) \tag{7}$$

Then $L_p(E)$ is the Fréchet derivative of $p(\cdot)$ at A in the matrix direction E.

Denote the Fréchet derivative of $p(\lambda, A)$ in the direction E as $L_p(\lambda; A, E)$.

Proposition 2: Assume that the system matrix in (1) or (6) is perturbed according to $A = A_0 + E$ and least-squares observer (5) for either case is evaluated for the nominal value of the state matrix A_0. Then, provided that $p(\lambda, \cdot)$ is still analytical at $\sigma(A)$ for all $\lambda \in \Lambda$, the observer estimation error $e(t) = \hat{x}(t) - x(t)$ is given by

$$\begin{aligned} e(t) &= \Big(\sum_{\lambda_i \in \Lambda} p(\lambda_i, A_0)^T C^T C \times \\ &\quad \big(L_p(\lambda_i; A_0, E) + O(\|E\|^2) I \big) \Big) x(t) \end{aligned} \tag{8}$$

Proof: Omitted. ∎

Proposition 2 justifies the following approximation

$$e(t) \approx \left(\sum_{\lambda_i \in \Lambda} p(\lambda_i, A_0)^T C^T C L_p(\lambda_i; A_0, E) \right) x(t) \tag{9}$$

Apparently, in order to desensitize the observer against system matrix uncertainty, the matrix function $p(\lambda_i, A_0)$ should be small, measured in some suitable norm, and satisfy certain smoothness conditions in the perturbation matrix direction E so that its Fréchet derivative is small, too.

IV. Evaluation of Fréchet derivatives

In this section, Fréchet derivatives of the matrix functions arising in least-squares state estimation are evaluated.

A. Continuous time operators

In the literature, three operators used for continuous least-squares observer design can be found, namely the differential operator, the continuous time-delay operator and a finite memory convolution operator, see [13],[7]. Though the differential operator cannot be used in practice for the observer implementation without some sort of additional filtering, it is taken into consideration for reference. Besides, the obtained results appear to be useful in the case of discrete-time shift operator.

An apparent difficulty in constructing operators for least-squares state estimation in continuous time is the necessity of finite memory, i.e. the operator impulse response has to be equal to zero starting from some finite time. In any other case, exact (deadbeat) state estimation is impossible. For instance, in [21], the integral operator $p_i(\lambda, s) = 1/s^\lambda$ is utilized for initial condition estimation. Since the integral operator possesses an infinite impulse response, any attempt to use it for least-squares estimation of the current value of the plant state vector would not be successful due to the reason above. However, choosing a finite integration horizon leads to a viable solution, [10].

The differential operator $p_d(\lambda, s) = s^\lambda$ where $\lambda = 1, 2, \ldots$ evaluated at a matrix A gives rise to $p_d(\lambda, A) = A^\lambda$. In [18], the Fréchet derivative of $p_d(A)$ is evaluated as

$$\begin{aligned} \lambda = 2: \quad L_{p_d}(A, E) &= AE + EA \\ \lambda \geq 2: \quad L_{p_d}(A, E) &= \sum_{\substack{j+k=\lambda-1 \\ j,k \geq 0}} A^j E A^k \end{aligned} \tag{10}$$

Another classical choice of implementation operator is the continuous time delay operator $p_\tau(\lambda, s) = \mathrm{e}^{\lambda s}$ where $\lambda < 0$. The corresponding matrix function is the matrix exponential $p_\tau(\lambda, A) = \mathrm{e}^{\lambda A}$ whose Fréchet derivative is [18]

$$L_{p_\tau}(A, E) = \int_0^1 \mathrm{e}^{(1-\theta)A\lambda} E \mathrm{e}^{\theta A \lambda} \, d\theta \tag{11}$$

An operator that combines integrating behavior with a finite memory is the finite-memory convolution operator

$$p_c(\lambda, s) = \frac{1 - \mathrm{e}^{(\lambda - \tau)s}}{s - \lambda} \tag{12}$$

Letting $\lambda = 0$ reduces $p_c(\lambda, \cdot)$ to finite-memory integration, utilized for instance in [10]. Being evaluated for a matrix argument, the finite-memory convolution operator yields

$$p_c(\lambda, A) = (I - \exp((\lambda - \tau)A))(A - \lambda I)^{-1} \quad (13)$$

The Fréchet derivative of p_c is not readily available and has to be derived separately.

Proposition 3: The Fréchet derivative of $p_c(\lambda, A)$ in the matrix direction E is given by

$$L_{p_c}(A, E) = \Big((\lambda - \tau)\int_0^1 \exp\big((1-\theta)(\lambda-\tau)A\big) E \exp(\theta(\lambda-\tau)A)\ d\theta + p_c(\lambda, A)E\Big)(\lambda I - A)^{-1}. \quad (14)$$

Proof: Omitted. ■

B. Discrete time operators

Using z^λ as a parametrization operator for (5) in discrete time will obviously lead exactly to the same form of the Fréchet derivative as for s^λ in continuous time.

Another popular discrete operator is the δ-operator, [22]

$$p_\delta(\lambda, z) = \left(\frac{z-1}{\Delta}\right)^\lambda$$

where Δ is a constant. δ-operator can be interpreted as a discrete time approximation of the continuous time derivative $\frac{d(\cdot)}{dt}$. The corresponding matrix function is

$$p_\delta(\lambda, A) = \left(\frac{A-I}{\Delta}\right)^\lambda$$

The Fréchet derivative of $p_\delta(1, z)$ is

$$L_{p_\delta}(1; A, E) = \frac{1}{\Delta}E$$

Recall now that the Fréchet derivative of z is simply E. Immediately one obtains that a parametrization in δ-operator is less sensitive to matrix perturbations than a parametrization of the same system in z whenever $\Delta > 1$. This is a well-known fact, see [17].

The use of so-called γ-operator

$$p_\gamma(\lambda, z) = \left(\frac{1}{\Delta}\frac{z-1}{z+1}\right)^\lambda$$

has been advocated in [17]. Assuming $\sigma(A) \cap -1 \equiv \emptyset$ and evaluating $p_\gamma(1, z)$ at $z = A$ gives

$$p_\gamma(A) = \frac{1}{\Delta}(A-I)(A+I)^{-1}$$

It can be shown that

$$L_{p_\gamma}(1; A, E) = \frac{2}{\Delta}(A+I)^{-1}E(A-I)$$

In contrast with z and δ-operator, the sensitivity of γ-operator depends on the properties of A, as also pointed out in [17]. It can be shown that for the same value of Δ, the sensitivity of γ-operator can be higher or lower than that of δ-operator, for different values of the plant system matrix.

A discrete counterpart to the continuous operator $p_c(\lambda, \cdot)$ given by (12) is the operator

$$p_{dc}(\lambda, z) = \frac{1 - (\lambda z^{-1})^\tau}{z - \lambda} = \sum_{i=0}^{\tau-1} \lambda^i z^{-(i+1)}$$

In contrast to the continuous finite-memory convolution, $p_{dc}(\lambda, z)$ can be written as a finite polynomial in z of order τ, e.g.

$$\begin{aligned} p_{dc}(\lambda, z)|_{\tau=1} &= z^{-1} \\ p_{dc}(\lambda, z)|_{\tau=2} &= z^{-1}(1 + \lambda z^{-1}) \end{aligned}$$

Being evaluated for a nonsingular matrix argument, the discrete finite-memory convolution operator gives

$$p_{dc}(\lambda, A) = (I - \lambda^\tau A^{-\tau})(A - \lambda I)^{-1} = \sum_{i=0}^{\tau-1} \lambda^i A^{-(i+1)}$$

Therefore, the Fréchet derivative of $p_{dc}(\lambda, A)$ can be obtained using the same rules as for the differential operator, see (10), e.g. for $\tau = 2$

$$L_{p_{dc}}(A, E) = -A^{-1}\left(\lambda(A^{-1}E + EA^{-1}) + E\right)A^{-1}$$

V. Least-squares state estimator of a harmonic oscillator

Consider now a single-tone harmonic oscillator, i.e. autonomous dynamic system

$$\begin{aligned} \dot{x}(t) &= Fx(t) \\ y(t) &= Cx(t) \end{aligned} \quad (15)$$

whose matrices are the following ones

$$\begin{aligned} F &= \begin{bmatrix} 0 & \omega^2 \\ -1 & 0 \end{bmatrix} \\ C &= \begin{bmatrix} 1 & 1 \end{bmatrix} \end{aligned} \quad (16)$$

The choice of this particular example is motivated by the ease of its generalization to the case of multiple frequencies as well as its simplicity and practical importance in engineering applications.

Proposition 4: Assume that for any $\lambda \in \Lambda$, the operator $(P\cdot)(\lambda; t)$ satisfies Assumption 1. Then, the observer

$$\hat{x}(t) = U\mathcal{V}_F^{-1}\sum_{\lambda\in\Lambda}\begin{bmatrix} p(\lambda, j\omega) \\ p(\lambda, -j\omega) \end{bmatrix}(Py)(\lambda; t) \quad (17)$$

where

$$\begin{aligned} U &= \begin{bmatrix} \frac{-j\omega}{1-j\omega} & \frac{j\omega}{1+j\omega} \\ \frac{1}{1-j\omega} & \frac{1}{1+j\omega} \end{bmatrix}; \\ \mathcal{V}_F &= \sum_{\lambda\in\Lambda}\begin{bmatrix} p^2(\lambda, j\omega) & p(\lambda, j\omega)p(\lambda, -j\omega) \\ p(\lambda, j\omega)p(\lambda, -j\omega) & p^2(\lambda, -j\omega) \end{bmatrix} \end{aligned}$$

yields the estimate $\hat{x}(t) = x(t), t > \tau$ if and only if there is no $c \in \mathrm{C}$ such that

$$p(\lambda, j\omega) = cp(\lambda, -j\omega), \forall \lambda \in \Lambda \quad (18)$$

Proof: See Theorem in [12] ■

Since only ω can vary in F, it is reasonable to choose the following perturbation direction

$$E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Proposition 5: The Fréchet derivative of $p(\lambda, F)$ in the direction E is

$$L_p(F,E) = \frac{1}{2} \sum_{s=\pm j\omega} p'(s,\lambda) \begin{bmatrix} 1 \\ -\frac{1}{s} \end{bmatrix} \begin{bmatrix} -\frac{1}{s} & 1 \end{bmatrix}$$

and its Euclidean norm $\|L_p(F,E)\|_2$ is

$$\|L_p(F,E)\|_2^2 = \frac{1}{2}\left(\left(1+\frac{1}{\omega^4}\right)\alpha^2(\omega,\lambda) + \frac{2}{\omega^2}\beta^2(\omega,\lambda) + |1-\frac{1}{\omega^2}||\beta(\omega,\lambda)|\sqrt{\left(\frac{1}{\omega^2}+1\right)^2\alpha^2(\omega,\lambda)+\frac{4}{\omega^2}\beta^2(\omega,\lambda)}\right)$$

where

$$p'(j\omega,\lambda) = \alpha(\omega,\lambda) + j\beta(\omega,\lambda)$$

Proof: Omitted. ■

A. Time delay operator

The time delay operator is the simplest one and widely used for implementation of least-squares state estimators. Given

$$p'_\tau(j\omega,\lambda) = -\lambda\left(\cos(\omega\lambda) - j\sin(\omega\lambda)\right)$$

the norm of the Fréchet derivative is

$$\|L_{p_\tau}(F,E)\|_2^2 = \frac{\lambda^2}{2}\left((1-\frac{1}{\omega^2})^2\cos^2(\omega\lambda) + \frac{2}{\omega^2} + |\frac{1}{\omega^2}-1||\cos^2(\omega\lambda)|\sqrt{(1-\frac{1}{\omega^2})^2\cos^2(\omega\lambda)+\frac{4}{\omega^2}}\right) \quad (19)$$

Clearly, the problem of global minimization of $\|L_{p_\tau}(F,E)\|_2$ with respect to the time delay duration λ has only a trivial and not feasible solution, i. e.

$$\arg\min_\lambda \|L_{p_\tau}(F,E)\|_2 = 0$$

Despite of that, (19) provides an important insight into how the time delay value influences the directional sensitivity of the matrix operator. Indeed, one can notice that

$$\|L_{p_\tau}(F,E)\|_2^2 = \frac{\lambda^2}{2} b_\tau(\lambda)$$

where $b_\tau(\lambda)$ is periodic and bounded from below and above. Besides, $b_\tau(\lambda)$ achieves minima when $\cos(\lambda\omega) = 0$ and maxima when $\cos^2(\lambda\omega) = 1$. This observation results in the following bounds, see Fig. 1

$$\frac{\lambda}{\omega} \le \|L_{p_\tau}(F,E)\|_2 \le \frac{\lambda}{\omega^2} \qquad \text{for } \omega < 1$$

$$\frac{\lambda}{\omega} \le \|L_{p_\tau}(F,E)\|_2 \le \lambda \qquad \text{for } \omega > 1$$

Therefore, to achieve low sensitivity to frequency variations, the time delays in the estimator and the quantities $\cos(\lambda\omega)$ should be kept as small as possible. As well known, small time delays in least-squares observers lead to drastic estimation error transients and should be avoided. Thus, the only option left here is to minimize $\cos(\lambda\omega)$.

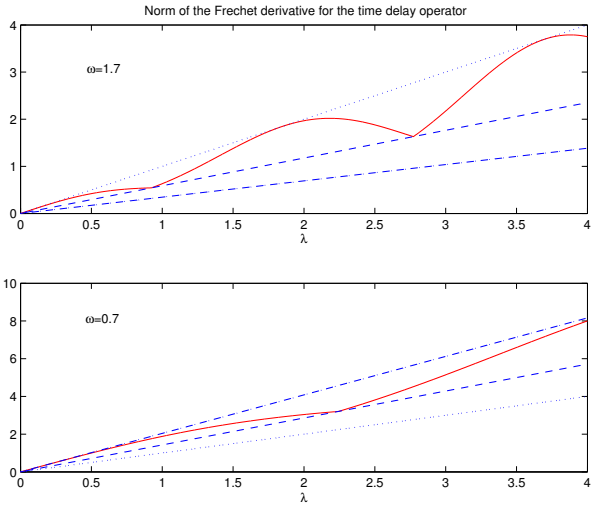

Fig. 1. $\|L_{p_\tau}(F,E)\|_2$ for the time delay operator and its bounds. Upper plot for $\omega > 1$, lower plot for $\omega < 1$. Solid line — $\|L_{p_\tau}(F,E)\|_2$. Dashed line — λ/ω. Dash-dotted line — λ/ω^2. Dotted line — λ.

B. Finite memory convolution operator

For $\tau > 0$, in this case

$$p(\lambda,s) = p_c(\lambda,s) = \frac{1-e^{(\lambda-s)\tau}}{s-\lambda}$$

Notice here that the sign of λ does not influence stability of $p(\lambda,s)$ due to the finite memory of it. Actually, it is easy to see that the pole of $p(\lambda,s)$ is canceled yielding an infinite series representation of the operator

$$p(\lambda,s) = \sum_{i=1}^{\infty} \frac{(\lambda-s)^{i-1}\tau^i}{i!}$$

Furthermore,

$$\text{Re } p(\lambda,j\omega) = \frac{1}{\lambda^2+\omega^2}\left(\omega e^{\lambda\tau}\sin(\omega\tau) + \lambda(e^{\lambda\tau}\cos(\omega\tau)-1)\right)$$

$$\text{Im } p(\lambda,j\omega) = \frac{1}{\lambda^2+\omega^2}\left(\omega(e^{\lambda\tau}\cos(\omega\tau)-1) - \lambda e^{\lambda\tau}\sin(\omega\tau)\right)$$

For the reason explained later, consider now only $\lambda < 0$. In order to obtain a satisfactory transient performance in the observer, τ should not be small. Therefore, $e^{\lambda\tau} << 1$ and the following approximations apply

$$\text{Re } p_c(\lambda,j\omega) \approx -\frac{\lambda}{\lambda^2+\omega^2}$$

$$\text{Im } p_c(\lambda,j\omega) \approx -\frac{\omega}{\lambda^2+\omega^2}$$

Clearly, the gain of $p(\lambda,\cdot)$ is small for high frequencies ω and large in absolute values parameters λ. The derivative of the operator symbol is

$$p'_c(\lambda,s) = \frac{(1-(\lambda-s)\tau)\,e^{(\lambda-s)\tau}-1}{(s-\lambda)^2}$$

Under the same approximation assumptions as before, it follows

$$\|L_{p_c}(F,E)\|_2^2 \approx \frac{1}{2(\omega^2+\lambda^2)^4}\left(\left(1+\frac{1}{\omega^4}\right)(\omega^2-\lambda^2)^2+2\lambda^2 - \omega\lambda|1-\frac{1}{\omega^2}|\sqrt{\left(1+\frac{1}{\omega^2}\right)^2(\omega^2-\lambda^2)^2+4\lambda^2}\right)$$

and the norm of the Fréchet derivative as well as the sensitivity of the corresponding matrix function can be brought down by choosing $\lambda << 0$, see Fig. 2. Notice also that $\|L_{p_c}(F,E)\|_2$ continues to rise for positive values of λ which fact explains why only $\lambda < 0$ have been considered in the first place. As mentioned before, stability of the observer is preserved even for $\lambda > 0$ thanks to the finite memory of the operator. A standard and $L_2[t, t-\tau]$-optimal solution to

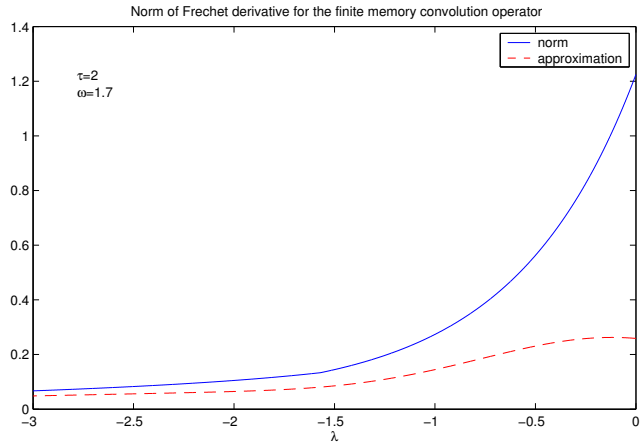

Fig. 2. $\|L_{p_c}(F,E)\|_2$ for the time finite memory convolution operator (solid line) and its approximation (dashed line). The approximation is better for $\lambda << 0$. Note that the approximation is not a bound.

finite-memory state estimation is to use

$$\hat{x}_H(t) = \int_{t-\tau}^{t} \exp\left(A^T(\theta-t)\right) C^T y(\theta)\, d\theta$$

where the pair (A, C) in (1) is chosen so that the corresponding gramian is balanced. Using Buchheim's formula [20], one has for n constant matrices M_{ij}

$$\exp\left(-A^T(t-\theta)\right) = \sum_j \sum_{\mu_i \in \sigma(A)} (t-\theta)^j e^{-\mu_i(t-\theta)} M_{ij}$$

Then, for the case of distinct eigenvalues of A, the optimal estimate $\hat{x}_H$ can be written in the form of (4) with the use of finite-memory convolution (12) and $\Lambda \equiv \sigma(-A)$. For a stable system, the latter implies positive values of λ in $p_c(\lambda, s)$ and results in high estimate sensitivity, as demonstrated in the analysis above.

VI. Conclusions

Sensitivity of least-squares state estimators to structured uncertainty in the system matrix of the plant is studied in a common for continuous and discrete case framework using the Fréchet derivative. The design degree of freedom offered by the estimator parametrization operator is investigated. It is shown that the state estimation error caused by the plant model mismatch is proportional to the Fréchet derivative of the symbol of the parametrization operator evaluated for the nominal value of the system matrix. The particular case of state estimation in a single-tone continuous oscillator is treated in detail as an illustration. When the time delay is used for observer implementation, the sensitivity of the observer in the direction of perturbation rises almost linearly with the time delay duration. On the contrary, the observer sensitivity decreases when a continuous finite-memory convolution operator with a significant decay rate of the kernel function is chosen for the design.

References

[1] James D. Gilchrist, "n-observability for linear systems," *IEEE Transactions on Automatic Control*, vol. 11, no. 3, pp. 388–395, 1966.

[2] Wook Hyun Kwon, Pyung Soo Kim, and Soo Hee Han, "A receding horizon unbiased FIR filter for discrete-time state space models," *Automatica*, vol. 38, pp. 545–551, 2002.

[3] Stanislav Fuksa and Witold Byrski, "General approach to linear optimal estimator of finite number of parameters," *IEEE Transactions on Automatic Control*, vol. 29, pp. 470, 1984.

[4] Wook Hyun Kwon and Oh Kyu Kwon, "FIR filters and recursive forms for continuous time-invariant state space models," *IEEE Transactions on Automatic Control*, vol. 32, no. 4, pp. 352–356, April 1987.

[5] A. Medvedev and G. Hillerström, "An external model control system," *In Control–Theory and Advanced Technology*, vol. 10, no. 4, Part 4, pp. 1643–1665, 1995.

[6] A. Medvedev, "State estimation and fault detection by a bank of continuous finite-memory filters," *International Journal of Control*, pp. 499–517, 1998.

[7] A. Medvedev, "Disturbance attenuation in finite spectrum assignment controllers," *Automatica*, vol. 33, no. 6, pp. 1163–1168, June 1997.

[8] C. V. Rao, J. B. Rawlings, and J. Lee, "Constrained linear state estimation - a moving horizon approach," *Automatica*, vol. 37, no. 10, pp. 1619–1628, 2001.

[9] Robert Engel and Gerhard Kreisselmeier, "A continuous-time observer that converges in finite time," vol. 47, no. 7, pp. 1202–1204, June 2002.

[10] Michel Fliess and Hebertt Sira-Ramirez, "State reconstructors: a possible alternative to asymptotic observers and Kalman filters," in *Proceedings of CESA*, Lille, (France), 2003.

[11] Johann Reger and Jerome Jouffroy, "On algebraic time-derivative estimation and deadbeat state reconstruction," Tech. Rep., http://www.reger-online.org/, September 2007.

[12] M. Bask and A. Medvedev, "Analysis of least-squares state estimators for a harmonic oscillator," in *IEEE Conference on Decision and Control, CDC2000*, Sydney, Australia, 2000, pp. 1800–1804.

[13] A. Medvedev, "Continuous least-squares observers with applications," *IEEE Transactions on Automatic Control*, vol. 41, pp. 1530–1536, Oct. 1996.

[14] Thomas Kailath, *Linear Systems*, 1980.

[15] D. H. Chyung, "State variable reconstruction," *International Journal of Control*, vol. 40, pp. 955–963, 1984.

[16] A. Medvedev and H. Toivonen, "Feedforward time–delay structures in state estimation. Finite memory smoothing and continuous deadbeat observers," *IEE Proceedings-D*, vol. 141, no. 2, pp. 121–129, 1994.

[17] Michel Gevers and Gang Li, *Parametrizations in Control, Estimation and Filtering Problems*, Springer-Verlag, 1993.

[18] Rajendra Bhatia, *Matrix Analysis*, New York Springer, 1997.

[19] Charles Kenney and Alan J. Laub, "Condition estimates for matrix functions," *SIAM J. Matrix Anal. Appl.*, vol. 10, no. 2, pp. 191–209, April 1989.

[20] C. A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[21] Z. Ding, "Change detection based on initial state estimation," *System & Control Letters*, vol. 21, pp. 337–345, 1993.

[22] Richard H. Middleton and Graham C. Goodwin, *Digital Control and Estimation: A Unified Approach*, Prentice Hall, Englewoods Cliffs, 1990.