Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 53-2 (2020) 7406–7411

Sequential processing and performance optimization in nonlinear state estimation

S. Battilotti *

* DIAG, Università "Sapienza" di Roma, Roma, Italy, (e-mail: battilotti@diag.uniroma1.it, stefano.battilotti@uniroma1.it).

Abstract: We propose a framework for designing observers for noisy nonlinear systems with global convergence properties and performing robustness and noise sensitivity. Our state observer is the result of the combination of a state norm estimator with a bank of Kalman-type filters, parametrized by the state norm estimator. The state estimate is sequentially processed through the bank of filters. In general, existing nonlinear state observers are responsible for estimation errors which are sensitive to model uncertainties and measurement noise, depending on the initial state conditions. Each Kalman-type filter of the bank contributes to improve the estimation error performances to a certain degree in terms of sensitivity with respect to noise and initial state conditions. A sequential processing algorithm for performance optimization is given and simulations show the effectiveness of these sequential filters.

Keywords: Noisy systems, nonlinear dynamics, Kalman-type filters and observers.

1. INTRODUCTION

Design of state observers for nonlinear systems with large domain of attraction, convergence speed performances, robustness and moderate sensitivity to (measurement) noise is a challenging task. Intrinsic limitations to this task have been studied for linear systems for instance in Seron et al. (1997). We have a vast literature on observer design for nonlinear systems (we will not reference this here for lack of space) focusing on single such performances, almost exclusively domain of attraction or convergence speed performances. In most of these contributions, robustness to model uncertainties and sensitivity to noise are not considered at all or considered only *a posteriori*, evaluating the possible effects on the error performances. An important conclusion of these works is that for high-gain observers (HGO) it is not possible to achieve large domains of attraction without downgrading the sensitivity to measurement noise. Only recently robustness to model uncertainties and sensitivity to noise have been taken into account in the design: for feedback linearizable systems in Khalil et al. (2014) and Prasov et al. (2013) and bounded state teajectories in Shim et al. (2016), where general Lyapunov-based conditions are given for observers with quasi-Disturbance-to-Error Stability (qDES). In Sanfelice et al. (2011) and in Battilotti (2009) the issue of somewhat reducing the sensitivity of global observers to additive measurement noise is discussed for some classes of nonlinear systems with bounded solutions. More recently, the work Astolfi et al. (2019) unites local observers, which have good error performances versus measurement noise like extended Kalman filters, with semiglobal HGO's, which have bad error performances. Systems with bounded solutions are considered and the resulting observer has a switching structure which guarantees the compromise between bad (semiglobal) and good (local) error performances but its correct working depends on some local and semiglobal norm estimators together with the exact knowledge of the domains of attraction of the local and semiglobal observers.

In this paper we want to consider a quite general class of nonlinear systems with model uncertainties (or state noise) and measurement noise and design global state observers with the primary objective to optimize the error performances in terms of robustness and noise sensitivity. The observer we propose in this paper consists of the following parts: I) a state norm estimator (SN) and II) a bank of sequential Kalman-type filters ($\mathrm{K}_1, \ldots, \mathrm{K}_p$) parametrized by the SN estimator. Each K-filter updates its gain matrix with the solution of a (differential) Riccati equation parametrized by the SN estimator and processes saturated estimates. One of the innovative feature of our observer relies in the bank of K-filters, which implements *de facto* a sequential optimization process of the error performances in terms of sensitivity versus noise and model uncertainties. Each filter K_j processes the state estimate $\hat{X}_{j-1}$, given by the previous filter K_{j-1}, by computing an estimate $\hat{x}_j$ of the relative displacement of the state from $\hat{X}_{j-1}$ and passing the new state estimate $\hat{X}_j := \hat{X}_{j-1} + \hat{x}_j$ to the next filter K_{j+1}. As shown by the simulations, the results are impressive and can be particularly appreciated with HGO's which have the worst error performances with respect to measurement noise.

2. NOTATION

(**N1**) $\mathbb{R}^n$ (resp. $\mathbb{R}^{n\times s}$) is the set of n-dimensional real column vectors (resp. $n \times s$ matrices). $\mathbb{R}_\geqslant$ (resp. $\mathbb{R}_>$) denotes the set of non-negative real numbers (resp. positive real numbers). $|a|$ denotes the absolute value of $a \in \mathbb{R}$, $\|a\|$ denotes the euclidean norm of $a \in \mathbb{R}^n$, $\|A\|$ denotes the norm of $A \in \mathbb{R}^{n\times n}$ induced from $\|\cdot\|$. Let $\mathcal{P}^n_>$ be the set of symmetric positive definite matrices $A \in \mathbb{R}^{n\times n}$.

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2020.12.1272

Moreover, $\sigma(A)$ denotes the spectrum of A. $\lambda^W_{min} := \min_{\lambda\in\sigma(W)} \mathbf{Re}\{\lambda\}$ and $\lambda^W_{max} := \max_{\lambda\in\sigma(W)}\{\lambda\}$, with $\mathbf{Re}$ denoting the real part.

(**N3**) Let $\mathcal{I}, \mathcal{J} \subset \mathbb{R}$ be intervals of the form $(a, +\infty)$ (resp. $[a, +\infty)$) and let $\mathcal{K}_\infty(\mathcal{I}, \mathcal{J})$ be the set of functions $\delta : \mathcal{I} \to \mathcal{J}$ such that $\lim_{s\to+\infty} \delta(s) = +\infty$ and $\frac{d\delta}{ds}(s) > 0$ for all $s > a$. By $Dom(\delta)$ (resp. $Dom(\delta^{-1})$) we denote the domain of δ (resp. δ^{-1}). A *saturation function* sat_l *with saturation levels* $l \in \mathbb{R}^n_>$ is a function $sat_l(x) := (sat_{l_1}(x_1), \ldots, sat_{l_n}(x_n))^T$, $x \in \mathbb{R}^n$, such that for each $i = 1, \ldots, n$ and $x_i \in \mathbb{R}$:

$$sat_{l_i}(x_i) = \begin{cases} x_i & x_i \in [-l_i, l_i], \\ \text{sign}(x_i) l_i & \text{otherwise}. \end{cases} \quad (1)$$

3. CLASS OF SYSTEMS

Consider a nonlinear system of the general form

$$\dot{x} = f(x, d), \; y = h(x, d) \quad (2)$$

with states $x \in \mathbb{R}^n$, outputs $y \in \mathbb{R}^p$, f and h continuously differentiable functions, and disturbances (or exogenous inputs) $d \in \mathcal{D}$, the space of bounded functions $d : [0, +\infty) \to D \subset \mathbb{R}^s$ with sup norm $\|d\|_\infty := \sup_{t\in[0,+\infty)} \|d_t\|$ uniformly bounded by a known $d_\infty > 0$ (more generally, it is possible to split d into a state disturbance d_s and a measurement disturbance d_y). We will denote $x(t), y(t)$ and $d(t)$ by $\mathbf{x}_t, \mathbf{y}_t$ and, respectively, $\mathbf{d}_t$. More conveniently, we may represent $\mathbf{x}_t$ (resp. $\mathbf{y}_t$) as $\mathbf{x}_t(x, s; d)$ (resp. $\mathbf{y}_t(x, s; d)$) to denote the value at time t of the (unique) solution (resp. output) of system (2) with input d and initialized at x at time s. Our problem is to design state estimators for (2) with good performances in terms of sensitivity to model uncertainties (represented by state disturbances) and measurement noise. All the results of this paper can be straightforwardly extended to systems (2) with inputs. In our study we discard finite exit times of the forward solutions of (2) from any bounded domain (this is the case after a change of time scale in (2) if, for instance, we assume suitable unboundedness observability conditions for (2)).

(A0). *(Forward completeness) The solutions* $\mathbf{x}_t(x, s; d)$ *of (2) are defined for all* $s \in \mathbb{R}_\geqslant$, $d \in \mathcal{D}$ *and* $(x, t) \in \mathbb{R}^n \times [s, +\infty)$.

4. STATE NORM ESTIMATORS

The first issue we want to discuss is the design of a state norm (SN) estimator for (2). Our assumption on (2) is:

(A1) *(SN estimator). There exist* $\lambda_v > 0$, $\lambda_h, \lambda_d \in \mathcal{K}_\infty(\mathbb{R}_\geqslant, \mathbb{R}_\geqslant)$ *and continuously differentiable* $v : \mathbb{R}^n \times \mathbb{R}_\geqslant \to \mathbb{R}$ *such that for all* $(x, t) \in \mathbb{R}^n \times \mathbb{R}_\geqslant$

$$(PDI) : \frac{\partial v}{\partial x}(x, t) f(x, \mathbf{d}_t) + \frac{\partial v}{\partial t}(x, t) \leqslant -\lambda_v v(x, t) + \lambda_h(\|h(x, \mathbf{d}_t)\|) + \lambda_d(d_\infty). \quad (3)$$

Moreover, there exist $\beta \in \mathcal{K}_\infty(\mathbb{R}_\geqslant, \mathbb{R})$, $\epsilon_0 \in \mathcal{K}_\infty(\mathbb{R}_\geqslant, \mathbb{R}_\geqslant)$ *and* $\epsilon_1, \bar{t} > 0$ *such that for all* $(x, t) \in \mathbb{R}^n \times [\bar{t}, +\infty)$:

$$v(x, t) \geqslant \beta(\|x\|), \quad (4)$$

$$\lambda_h(\|h(x, \mathbf{d}_t)\|) \leqslant \epsilon_1 \beta(\|x\|) + \epsilon_0. \quad (5)$$

As in Kirchman et al. (2001), the interest in the PDI (3) is motivated by the following result which establishes the existence of a SN estimator for (2). We assume without loss of generality that $\beta(0) \leqslant 0$ in such a way that $Dom(\beta^{-1}) \supseteq \mathbb{R}_\geqslant$.

Proposition 1. Assume there exist a continuously differentiable function $v : \mathbb{R}^n \times \mathbb{R}_\geqslant \to \mathbb{R}$, *together with* $\lambda_v > 0$, $\lambda_d, \lambda_h \in \mathcal{K}_\infty(\mathbb{R}_\geqslant, \mathbb{R}_\geqslant)$, $\beta \in \mathcal{K}_\infty(\mathbb{R}_\geqslant, \mathbb{R})$ *and* $\bar{t} > 0$ *such that (3) and (4) hold true for all* $(x, t) \in \mathbb{R}^n \times \mathbb{R}_\geqslant$ *and, respectively, for all* $(x, t) \in \mathbb{R}^n \times [\bar{t}, +\infty)$. *For each* $d \in \mathcal{D}$, $x_0 \in \mathbb{R}^n$, $\hat{v}_0 \in \mathbb{R}$ *and* $\hat{\gamma} > 0$ *there exists* $\bar{t}_{x_0,d} \geqslant \bar{t}$ *such that for* $t \geqslant \bar{t}_{x_0,d}$:

$$\|\mathbf{x}_t(x_0, 0; d)\| \leqslant \beta^{-1}(\hat{\mathbf{v}}_t(\hat{v}_0, 0; y) + \hat{\gamma}),$$

where $\hat{\mathbf{v}}_t(\hat{v}_0, 0; y)$ *is the solution of*

$$\dot{\hat{\mathbf{v}}}_t = -\lambda_v \hat{\mathbf{v}}_t + \lambda_h(\|\mathbf{y}_t\|) + \lambda_d(d_\infty), \hat{v}_0 \in \mathbb{R}. \quad (6)$$

Therefore, having (3), (4) at hand makes available an (exponentially) converging SN estimator (6). We refer the reader to the more detailed Battilotti (2019) for many constructive conditions for solutions of the PDI (3) and the uniform lower bound (4).

Remark 1. The additional condition (5) requires that the growth of $v(x, t)$ along the system's trajectories is at most exponential. Moreover, (4)-(5) imply that $\lambda_h(\|h(x, \mathbf{d}_t)\|) \leqslant \epsilon_1 v(x, t) + \epsilon_0$, which is a reasonable condition in view of the fact that, under a suitable backward completeness assumption on (2), a solution of the PDI (3) is $v(x, t) = \int_0^t e^{-\lambda_v(t-s)} \lambda_h(\|h(\mathbf{x}_s(x, t; d), \mathbf{d}_s)\|) ds$ and for $t \geqslant \bar{t} > 0$

$$v(x, t) \geqslant e^{-\lambda_v \bar{t}} \int_0^{\bar{t}} \lambda_h(\|h(\mathbf{x}_s(x, \bar{t}; \bar{d}), \bar{\mathbf{d}}_s)\|) ds$$

with

$$\lim_{\bar{t}\to 0^+} \sup_{x\in\mathcal{X}} \left| \frac{1}{\bar{t}} \int_0^{\bar{t}} \lambda_h(\|h(\mathbf{x}_s(x, \bar{t}; \bar{d}), \bar{\mathbf{d}}_s)\|) ds - \lambda_h(\|h(x, \mathbf{d}_t)\|) \right| = 0$$

for any compact $\mathcal{X} \subset \mathbb{R}_n$ and $\bar{\mathbf{d}}_s := \mathbf{d}_{s+t-\bar{t}}$ (see Battilotti (2019)). ◁

5. K-FILTERS PARAMETRIZED BY STATE NORM ESTIMATORS

The second task of our work is to identify canonical classes of observers (K-filters) for (2) under quite general assumptions and parametrized by the SN estimator of the previous section. These K-filters will be sequentially connected for performance optimization and, together with the SN estimator, form the state observer which we propose in this paper. To this aim, we consider a system (2) in which the linear part is highlighted:

$$\dot{x} = f(x, d) := Ax + f_0(x, d),$$
$$y = h(x, d) := Cx + h_0(x, d) \quad (7)$$

with $\frac{\partial f_0}{\partial x}(0, d) = 0$ and $\frac{\partial h_0}{\partial x}(0, d) = 0$ for all d. In order to identify a canonical structure for our K-filters, we require the following features: the observer gain matrix is updated according to the solution of a (differential) Riccati equation, parametrized by a SN estimator, and the estimates processed by the observer are saturated with saturation levels parametrized by the same SN estimator. This parametrization is done through the action of certain generalized *families of dilations*, which is one of the innovative feature of our observers. Let $\Lambda : [1, +\infty) \to \mathbb{R}^{n\times n}$ be the continuously differentiable solution of the following matrix differential equation

$$\frac{d\Lambda}{ds}(s) = \frac{W\Lambda(s)}{s},\ \Lambda(1) = I, \tag{8}$$

where $W \in \mathbb{R}^{n\times n}$ is any diagonal matrix having all its eigenvalues in $\mathbb{C}^-$ (W being diagonal is assumed for simplicity and can be relaxed). As well known, this solution is $\Lambda(s) := e^{\ln(s)W}$, $s \geqslant 1$. We call Λ^{-1} (the inverse of Λ) a *family of dilations* if Λ is the solution of a matrix differential equation of the form (8). Notice that since $W = \mathrm{diag}\{W_{11}, \ldots, W_{nn}\}$ then $\Lambda(s) = \mathrm{diag}\{s^{W_{11}}, \ldots, s^{W_{nn}}\}$ but we retain the general notation $\Lambda(s) = e^{\ln(s)W}$ for the sake of generality (i.e. non diagonal W).

5.1 Discussion of the main assumptions on (7)

The notion of family of dilations is naturally associated with the notion of (generalized) homogeneity in the ∞-limit.

(A2) *(Generalized homogeneous linearization in the ∞-limit). There exist a family of dilations Λ^{-1}, degrees $\gamma_A, \gamma_C : [1, +\infty) \to [1, +\infty)$ and observable pair (C_0, A_0) such that*

$$\lim_{s\to+\infty} \frac{\Lambda(s)A\Lambda^{-1}(s)}{\gamma_A(s)} = A_0,\ \lim_{s\to+\infty} \frac{C\Lambda^{-1}(s)}{\gamma_C(s)} = C_0.$$

Next, we introduce similar assumptions (in the upper bound) to dominate the nonlinearities via the homogeneous linearization.

(A3) *(Generalized homogeneous domination). For all $x \in \mathbb{R}^n$, $d \in D$ and $s \geqslant 1$*

$$\frac{1}{\gamma_A(s)}\Big\|\Lambda(s)\frac{\partial f_0(w,0)}{\partial w}\Big|_{w=\Lambda^{-1}(s)x}\Lambda^{-1}(s)\Big\| \leqslant F_0(x)$$
$$\frac{1}{\gamma_C(s)}\Big\|\frac{\partial h_0(w,0)}{\partial w}\Big|_{w=\Lambda^{-1}(s)x}\Lambda^{-1}(s)\Big\| \leqslant H_0(x)$$
$$\frac{1}{\gamma_A(s)}\Big\|\Lambda(s)\frac{\partial f_0(w,d)}{\partial d}\Big|_{w=\Lambda^{-1}(s)x}\Big\| \leqslant \gamma_d(s)F_d(x,d)$$
$$\frac{1}{\gamma_C(s)}\Big\|\frac{\partial h_0(w,d)}{\partial d}\Big|_{w=\Lambda^{-1}(s)x}\Big\| \leqslant \gamma_d(s)H_d(x,d)$$

where $\gamma_d : [1, +\infty) \to \mathbb{R}_>$ and $F_0(\cdot), H_0(\cdot)$, $F_d(\cdot,\cdot)$ and $H_d(\cdot,\cdot)$ are continuous nonnegative functions such that $F_0(0) = 0$ and $H_0(0) = 0$.

Remark 2. If f_0 and h_0 are globally Lipschitz with observable (C, A) **(A2)** and **(A3)** are satisfied with the simple choice $W = -I$, $\gamma_A(s) = 1$, $\gamma_C(s) = s$ and $\gamma_d(s) = 1/s$. ◁

Next, some condition on the asymptotic behavior of the function γ_d in **(D2)** is introduced, i.e. γ_d is a sufficiently fast decreasing function.

(A4) *(Asymptotic behavior of γ_d). There exist $\nu_0, \nu_1 > 0$ such that*

$$\Big|\frac{d\gamma_d}{ds}(s)\Big| \leqslant \nu_0\frac{\gamma_d(s)}{s} \leqslant \nu_1 s^{-(|\lambda^W_{min}|+1)},\ \forall s \geqslant 1. \tag{9}$$

Finally, we require a stability margin condition for error convergence.

(A5) *(Stability margin). For each $\lambda_\pi > 0$ there exist $\Pi \in \mathcal{P}^n_>$ and $c, \lambda_f > 0$ such that*

$$\begin{aligned}&\mathbf{Ric}(c, \lambda_\pi, \lambda_f, \Pi, W) \\ &\quad := \Pi(A_0 + \phi W + \lambda_\pi I) + (A_0 + \phi W + \lambda_\pi I)^\top\Pi \\ &\quad -C_0^\top C_0 + \frac{1}{\lambda_f}\Pi^2 + 2\lambda_f \sup_{\substack{|\mathcal{X}_i|\leqslant 2c,\\ i=1,\ldots,n}} \Big(F_0^2(\mathcal{X}) + H_0^2(\mathcal{X})\Big) \leqslant 0\end{aligned} \tag{10}$$

for all

$$\phi \in \Big[-\frac{\lambda_v}{|\lambda^W_{max}|}, \frac{\epsilon_1}{|\lambda^W_{max}|}\Big(1 + \frac{\epsilon_0(d_\infty) + \lambda_d(d_\infty)}{\epsilon_1}\Big)\Big].$$

Remark 3. Formula (10) is a Riccati inequality with unknown Π and all bounded matrices. Notice that since $\overline{\mathcal{F}}_0(0) = 0, \overline{\mathcal{H}}_0(0) = 0$, the contribution of the last two terms in (10) can be made arbitrarily small by selecting c arbitrarily small and λ_f arbitrarily large. For these reasons, it is sufficient to solve the inequality (10) by omitting the last two terms. On account of this, existence of solutions $\Pi \in \mathcal{P}^n_>$ for (10) is guaranteed for instance when (C_0, A_0) is in canonical observer form and W is diagonal or, alternatively, under detectability assumptions on $(C_0, A_0 + \phi W + \lambda_\pi I)$. ◁

5.2 The canonical K-filter

Under the given assumptions, a K-filter for (7) is

$$\begin{aligned}\dot{\hat{\mathbf{x}}}_t &= A\hat{\mathbf{x}}_t + \hat{f}_0(\hat{\mathbf{x}}_t, \hat{z}(\hat{\mathbf{v}}_t)) + K(\hat{z}(\hat{\mathbf{v}}_t))(\mathbf{y}_t - \hat{\mathbf{y}}_t),\\ \hat{\mathbf{y}}_t &:= C\hat{\mathbf{x}}_t + \hat{h}_0(\hat{\mathbf{x}}_t, \hat{z}(\hat{\mathbf{v}}_t)),\end{aligned} \tag{11}$$

$$\dot{\hat{\mathbf{v}}}_t = -\lambda_v\hat{\mathbf{v}}_t + \lambda_h(\|\mathbf{y}_t\|) + \lambda_d(d_\infty),\ \hat{\mathbf{v}}_0 > 0, \tag{12}$$

with

$$K(\hat{z}) := \frac{\gamma_A(\hat{z})}{\gamma_C^2(\hat{z})}P^{-1}(\hat{z})C^T,\ P(\hat{z}) := \Lambda^T(\hat{z})\Pi\Lambda(\hat{z}),$$

$$\begin{aligned}\hat{f}_0(\hat{x}, \hat{z}) &:= f_0\Big(\Lambda^{-1}(\hat{z})sat_c(\Lambda(\hat{z})\hat{x}), 0\Big),\\ \hat{h}_0(\hat{x}, \hat{z}) &:= h_0\Big(\Lambda^{-1}(\hat{z})sat_c(\Lambda(\hat{z})\hat{x}), 0\Big),\end{aligned} \tag{13}$$

for some $\hat{z} \in \mathcal{K}_\infty([1, +\infty), [1, +\infty))$ and $c > 0$ (to be specified from the design). Notice how the SN estimator (12) updates, through the action of the family of dilations Λ^{-1}, both the gain matrix K and the saturated estimates in the K-filter (11). Our main result for the single K-filter (11) is the following.

Theorem 1. Assume **(A0)**-**(A5)**. *There exist $\hat{z} \in \mathcal{K}_\infty$ $([1, +\infty), [1, +\infty))$, $\lambda_L > 0$ and $c > 0$ such that along the solutions of (7), (11) and for all $\mathbf{x}_0 \in \mathbb{R}^n$*

$$\limsup_{t\to+\infty} \|\mathbf{x}_t - \hat{\mathbf{x}}_t\| \leqslant \rho(d_\infty, c)d_\infty, \tag{14}$$

where

$$\rho(s,c) := \sqrt{\frac{2}{\lambda_{min}(\Pi)\lambda_L}}\frac{\Phi(s,c)\nu_1}{\nu_2}, \quad (15)$$
$$\Phi(s,c) := \sup_{\substack{|r|\leqslant s \\ |w_i|\leqslant c \\ i=1,\ldots,n}} (F_d(w,r) + H_d(w,r)).$$

Remark 4. The asymptotic bound (14) has many implications. For example, the noise sensitivity in (14) can be reduced to any degree in the case $F_d(0,d) = 0$ and $H_d(0,d) = 0$ for all d (multiplicative state and measurement disturbances) by choosing c (the saturation level of the saturated estimates) smaller and smaller and therefore reducing the magnitude of the term $\Phi(d_\infty, c)$ in (14). On the other hand, if $H_d(0,d) \neq 0$ (additive measurement disturbances) this kind of noise sensitivity reduction is possible only up to a certain degree. Moreover, the asymptotic bound (14) does not depend on the initial state x_0 or the magnitude of $\mathbf{x}_t$ or its bound $\widehat{\mathbf{z}}_t$ (as it happens mostly: see for instance Battilotti (2009), Sanfelice et al. (2011)) but only on the disturbance magnitude d_∞. Therefore, the error performances (in term of robustness and noise sensitivity) of the K-filter (11) are good even for large state initial conditions. ◁

5.3 Sensitivity to measurement noise

The achievement of the asymptotic bound (14), and consequently also of good performances in terms of sensitivity to measurement noise, strongly depends on the compatibility between the last conditions in **(A2)** and **(A3)**, on one hand, and the right inequality in **(A4)**, on the other, which may not hold at the same time for example when $\lim_{s\to+\infty} \gamma_C(s)s^{-|\lambda^W_{min}|} = 0$. This happens, for instance, for typical HGO's and feedback linearizable systems (Khalil et al. (2014), Andrieu et al. (2017), Prasov et al. (2013), Khalil (2017), Esfandiari et al. (2019)) where the matrix W of the family of dilations Λ^{-1} has monotonically increasing real eigenvalues: for this reason the function $\gamma_C(s)$ (which satisfies **(A2)** and **(A3)**) may have a slower behavior at infinity in comparison with $s^{|\lambda^W_{min}|}$. As a result, in these cases the estimation error bound is very sensitive to measurement noise. The aim of the next section is to optimize this bound through sequential K-filters by reducing the asymptotic gap between the functions $\gamma_C(s)$ and $s^{-|\lambda^W_{min}|}$. On the other hand, this reduction corresponds to reduce uniformly the gap among the (real parts of the) eigenvalues of W. As a matter of fact, the ideal situation is the one for which the eigenvalues of W are all equal and real, a typical situation when f_0 and h_0 are globally Lipschitz (see remark 2).

6. SEQUENTIAL K-FILTERS FOR ERROR BOUND PERFORMANCES OPTIMIZATION

When one single estimator (11) shows bad performances in terms of sensitivity to noise for the above mentioned reasons, in order to improve performances we sequentially process the estimate, given by (11), using multiple K-filters (11). In what follows, we denote $\widehat{z}(\widehat{\mathbf{v}}_t)$ simply by $\widehat{\mathbf{z}}_t$. By *perfomance optimization via sequential processing* we mean that if $\widehat{\mathbf{X}}_t^{(k)}$ is some estimate of $\mathbf{x}_t$ for which the estimation error $\widehat{\mathbf{e}}_t^{(k)} := \mathbf{x}_t - \widehat{\mathbf{X}}_t^{(k)}$ is (asymptotically) norm bounded by some $\Omega^{(k)}(\widehat{\mathbf{z}}_t)$, then an estimate $\widehat{\mathbf{x}}_t^{(k+1)}$ of $\widehat{\mathbf{e}}_t^{(k)}$ is designed by processing $\widehat{\mathbf{X}}_t^{(k)}$ in such a way that the error $\widehat{\mathbf{e}}_t^{(k+1)} := \widehat{\mathbf{e}}_t^{(k)} - \widehat{\mathbf{x}}_t^{(k+1)}$ is (asymptotically) norm bounded by some $\Omega^{(k+1)}(\widehat{\mathbf{z}}_t)$ which shows nicer properties than $\Omega^{(k)}(\widehat{\mathbf{z}}_t)$, say $\limsup_{t\to+\infty} \frac{\Omega^{(k+1)}(\widehat{\mathbf{z}}_t)}{\Omega^{(k)}(\widehat{\mathbf{z}}_t)} < 1$. In this sense, the state estimate $\widehat{\mathbf{X}}_t^{(k+1)} := \widehat{\mathbf{x}}_t^{(k+1)} + \widehat{\mathbf{X}}_t^{(k)}$ has better error performances than the state estimate $\widehat{\mathbf{X}}_t^{(k)}$ and the processing is repeated until significant improvements can be obtained.

Definition 1. If $\widehat{\mathbf{X}}_t^{(k)}$ and $\widehat{\mathbf{X}}_t^{(k+1)}$ are two state estimates such that:

$$\limsup_{t\to+\infty} \frac{\|\mathbf{x}_t - \widehat{\mathbf{X}}_t^{(k)}\|}{\Omega^{(k)}(\widehat{\mathbf{z}}_t)} = 0, \ \limsup_{t\to+\infty} \frac{\|\mathbf{x}_t - \widehat{\mathbf{X}}_t^{(k+1)}\|}{\Omega^{(k+1)}(\widehat{\mathbf{z}}_t)} = 0, \quad (16)$$

then we say that $\widehat{\mathbf{X}}_t^{(k+1)}$ outperforms $\widehat{\mathbf{X}}_t^{(k)}$ by $(1-\rho)$ $(\times 100)$ % if $\limsup_{t\to+\infty} \frac{\Omega^{(k+1)}(\widehat{\mathbf{z}}_t)}{\Omega^{(k)}(\widehat{\mathbf{z}}_t)} := \rho < 1$.

By iterating this sequential process for a certain number of times, we optimize the estimation error performances. This corresponds to the design of a bank of sequential K-filters : the k-th filter is responsible for the state estimate $\widehat{\mathbf{X}}_t^{(k)}$ which is sequentially processed by the $(k+1)$-th filter for computing the outperforming state estimate $\widehat{\mathbf{X}}_t^{(k+1)}$.

6.1 A recursive algorithm

In this section we give a recursive algorithm for sequential processing in the specific situation for which **(A4)** cannot be satisfied together with **(A2)** and **(A3)** (bad error performances) and state a somewhat general result for error bound optimization. This algorithm corresponds to design a suitable bank of canonical K-filters, each one processing the estimate given by the previous one.

Initialization step. Let

$$c^{(-1)} := c, \ c^{(0)} := \Phi(d_\infty, c)\sqrt{\frac{2}{\lambda_{min}(\Pi)}},$$
$$\Pi^{(0)} := \Pi, \ W^{(0)} := W, \ \Lambda^{(0)}(s) := \Lambda(s),$$
$$\gamma_A^{(0)}(s) := \gamma_A(s), \ \gamma_C^{(0)}(s) := \gamma_C(s), \ \gamma_d^{(0)}(s) := \gamma_d(s),$$

with Φ as in (15), Π as in **(A5)** and $\Lambda(s), \gamma_A(s), \gamma_C(s)$ and $\gamma_d(s)$ are as in **(A2)** and **(A3)**. Moreover,

$$\mathbf{e}_t^{(-1)} := \mathbf{x}_t, \ \widehat{\mathbf{e}}_t^{(-1)} := \widehat{\mathbf{x}}_t, \ \mathbf{e}_t^{(0)} := \mathbf{e}_t = \mathbf{e}_t^{(-1)} - \widehat{\mathbf{e}}_t^{(-1)},$$

where $\widehat{\mathbf{x}}_t$ is the estimate of $\mathbf{x}_t$, obtained as pointed out in theorem 1, and $\mathbf{e}_t$ the state estimation error. From the proof of theorem 1 (see Battilotti (2019) for details), we have

$$\limsup_{t\to+\infty} \|\Lambda^{(0)}(\zeta_t)\mathbf{e}_t^{(-1)}\| \leqslant c, \quad (17)$$

$$\limsup_{t\to+\infty} \frac{\|\Lambda^{(0)}(\zeta_t)\mathbf{e}_t^{(0)}\|}{\gamma_d^{(0)}(\zeta_t)} \leqslant c^{(0)}. \quad (18)$$

Let

$$\sigma(W^{(0)}) = \{\lambda_0^{(0)}, (1+h_0)\lambda_0^{(0)}, \ldots, (1+h_0(n-1))\lambda_0^{(0)}\}$$

for some $\lambda_0^{(0)} < 0, h_0 > 0$ and $\gamma_d^{(0)}(s) = s^{-k_d^{(0)}}$ for some $k_d^{(0)} \in (0, |\lambda_{min}^{W^{(0)}}|)$ (i.e. condition **(A4)** fails).

Step $k \geqslant 0$. Consider the error system

$$\dot{\mathbf{e}}_t^{(k)} = \Psi(\hat{\mathbf{e}}_t^{(k-1)}, \dot{\hat{\mathbf{e}}}_t^{(k-1)}) + A\mathbf{e}_t^{(k)} + f_0^{(k)}(\hat{\mathbf{e}}_t^{(k-1)}, \mathbf{e}_t^{(k)}, \mathbf{d}_t),$$
$$\mathbf{y}_t^{(k)} = C\mathbf{e}_t^{(k)} + h_0^{(k)}(\hat{\mathbf{e}}_t^{(k-1)}, \mathbf{e}_t^{(k)}, \mathbf{d}_t), \tag{19}$$

with

$$\mathbf{e}_t^{(k)} := \mathbf{e}_t^{(k-1)} - \hat{\mathbf{e}}_t^{(k-1)},\ \mathbf{y}_t^{(k)} := \mathbf{y}_t - C\hat{\mathbf{e}}_t^{(k-1)},$$
$$f_0^{(k)}(R, S, D) := f_0(R + S, D),$$
$$h_0^{(k)}(R, S, D) = h_0(R + S, D) \tag{20}$$

and let

$$\limsup_{t\to+\infty} \|\overline{\Lambda}^{(k)^{-1}}(\zeta_t)\Lambda^{(k)}(\zeta_t)\mathbf{e}_t^{(k-1)}\| \leqslant c^{(k-1)} \tag{21}$$

with $\overline{\Lambda}^{(k)}(s) := e^{\ln(s)\overline{W}^{(k)}}$ and

$$\sigma(\overline{W}^{(k)}) = \{\overline{\lambda}_0^{(k)}, (1 + h_0)\overline{\lambda}_0^{(k)}, \ldots, (1 + h_0(n - 1))\overline{\lambda}_0^{(k)}\}$$

for some $\overline{\lambda}_0^{(k)} \in [-\frac{k_d^{(k)}}{1+h_0(n-1)}, 0)$. Define the new family of dilations

$$\Lambda^{(k+1)}(s) := e^{\ln(s)W^{(k+1)}} = \overline{\Lambda}^{(k)^{-1}}(s)\Lambda^{(k)}(s)$$

with

$$\sigma(W^{(k+1)})$$
$$= \{\lambda_0^{(k+1)}, (1 + h_0)\lambda_0^{(k+1)}, \ldots, (1 + h_0(n - 1))\lambda_0^{(k+1)}\}$$

where $\lambda_0^{(k+1)} := \lambda_0^{(k)} - \overline{\lambda}_0^{(k)}$. Define the new functions

$$\gamma_A^{(k+1)}(s) := \left(\gamma_A^{(k)}(s)\right)^{\frac{\lambda_0^{(k+1)}}{\lambda_0^{(k)}}}, \gamma_C^{(k+1)}(s) := \left(\gamma_C^{(k)}(s)\right)^{\frac{\lambda_0^{(k+1)}}{\lambda_0^{(k)}}},$$
$$\gamma_d^{(k+1)}(s) := \left(\gamma_d^{(k)}(s)\right)^{\frac{\lambda_0^{(k+1)}}{\lambda_0^{(k)}}}. \tag{22}$$

Let $\hat{\mathbf{e}}_t^{(k)}$ be the estimate of $\mathbf{e}_t^{(k)}$ in (19), obtained with the canonical K-filter

$$\dot{\hat{\mathbf{e}}}_t^{(k)} = \Psi(\hat{\mathbf{e}}_t^{(k-1)}, \dot{\hat{\mathbf{e}}}_t^{(k-1)}) + A\hat{\mathbf{e}}_t^{(k)}$$
$$+ \hat{f}_0^{(k)}(\hat{\mathbf{e}}_t^{(k-1)}, \hat{\mathbf{e}}_t^{(k)}, \zeta_t) + K^{(k+1)}(\zeta_t)(\mathbf{y}_t^{(k)} - \hat{\mathbf{y}}_t^{(k)}),$$
$$\hat{\mathbf{y}}_t^{(k)} = C\hat{\mathbf{e}}_t^{(k)} + \hat{h}_0^{(k)}(\hat{\mathbf{e}}_t^{(k-1)}, \hat{\mathbf{e}}_t^{(k)}, \zeta_t), \tag{23}$$

where

$$K^{(k+1)}(Z) := \frac{\gamma_A^{(k+1)}(Z)}{\gamma_C^{(k+1)^2}(Z)} P^{(k+1)^{-1}}(Z)C^\top$$
$$P^{(k+1)}(Z) := \Lambda^{(k+1)^\top}(Z)\Pi^{(k+1)}\Lambda^{(k+1)}(Z),$$
$$\hat{f}_0^{(k)}(R, S, Z) := f_0^{(k)}(\Lambda^{(k+1)^{-1}} sat_{c^{(k)}+c^{(k-1)}}\{\Lambda^{(k+1)}(Z)R\},$$
$$\Lambda^{(k+1)^{-1}}(Z) sat_{c^{(k)}}\{\Lambda^{(k+1)}(Z)S\}, 0),$$
$$\hat{h}_0^{(k)}(R, S, Z) := h_0^{(k)}(\Lambda^{(k+1)^{-1}} sat_{c^{(k)}+c^{(k-1)}}\{\Lambda^{(k+1)}(Z)R\},$$
$$\Lambda^{(k+1)^{-1}}(Z) sat_{c^{(k)}}\{\Lambda^{(k)}(Z)S\}, 0) \tag{24}$$

and $\Pi^{(k+1)} \in \mathcal{P}_>^n$ and $\lambda_\pi^{(k+1)}, \lambda_f^{(k+1)} > 0$ be such that

$$\mathbf{Ric}\Big(\frac{3c^{(k)} + c^{(k-1)}}{2}, \lambda_\pi^{(k+1)}, \lambda_f^{(k+1)}, \Pi^{(k+1)}, W^{(k+1)}\Big) \leqslant 0. \tag{25}$$

Set

$$c^{(k+1)} := \Phi(d_\infty, 2c^{(k)} + c^{(k-1)})\sqrt{\frac{2}{\lambda_{min}(\Pi^{(k+1)})}} \tag{26}$$

and, finally, set

$$\hat{\mathbf{x}}_t^{(k)} := \hat{\mathbf{x}}_t^{(k-1)} + \hat{\mathbf{e}}_t^{(k)},$$
$$\mathbf{e}_t^{(k+1)} := \mathbf{e}_t^{(k)} - \hat{\mathbf{e}}_t^{(k)} = \mathbf{x}_t - \hat{\mathbf{x}}_t^{(k)}$$

and $k_d^{(k+1)} := \frac{\lambda_0^{(k+1)}}{\lambda_0^{(k)}} k_d^{(k)}$. From the proof of theorem 1

$$\limsup_{t\to+\infty} \frac{\|\Lambda^{(k+1)}(\zeta_t)\mathbf{e}_t^{(k+1)}\|}{\gamma_d^{(k+1)}(\zeta_t)} \leqslant c^{(k+1)}. \tag{27}$$

Goto step $k + 1$.

Remark 5. Each step k of the above algorithm can be completed upon (21) being satisfied. At each step k the stronger condition

$$\limsup_{t\to+\infty} \|\Lambda^{(k)}(\hat{\mathbf{z}}_t)\mathbf{e}_t^{(k-1)}\| \leqslant c^{(k-1)} \tag{28}$$

is guaranteed by the previous step $k - 1$. However, by increasing $\overline{\lambda}_0^{(k)} \in [-\frac{k_d^{(k)}}{1+h_0(n-1)}, 0)$ it is reasonably expected that, on account of (28), also (21) becomes satisfied. We can also try a larger $\overline{c}^{(k-1)} \geqslant c^{(k-1)}$ for which (21) becomes satisfied. In this case, we desaturate the estimate $\hat{\mathbf{e}}_t^{(k-1)}$ in $F_0^{(k)}$ and $H_0^{(k)}$ (see (24)) and let $\Pi^{(k+1)}(t)$ be the solution of a suitable differential Riccati equation (a suitable modification of **Ric** in (25)) parametrized by $\hat{\mathbf{e}}_t^{(k-1)}$. A modified algorithm overcoming this problem is under study. ◅

The main result of this section on error performances optimization is based on the above recursive algorithm.

Proposition 2. Assume **(A0)-(A3)** *and* **(A5)**. *For each* $k \geqslant 0$ *if*

$$\limsup_{t\to+\infty} \frac{1}{\hat{\mathbf{z}}_t^{\Delta^{(k+1)}}} \frac{c^{(k+1)}}{c^{(k)}} := \rho^{(k+1)} < 1, \tag{29}$$

$$\Delta^{(k+1)} := (1 - \frac{|\lambda_{min}^{W^{(k+1)}}|}{|\lambda_{min}^{W^{(k)}}|})(|\lambda_{min}^{W^{(k)}}| - k_d^{(k)}), \tag{30}$$

the state estimate $\hat{\mathbf{x}}_t^{(k+1)}$ *outperforms the state estimate* $\hat{\mathbf{x}}_t^{(k)}$ *by* $(1 - \rho^{(k+1)})$ *(* $\times 100$*) %.*

Remark 6. Notice that $0 < \Delta^{(k)}$ since $|\lambda_{min}^{W^{(k)}}| > |\lambda_{min}^{W^{(k+1)}}|$ and $k_d^{(k)} < |\lambda_{min}^{W^{(k)}}|$. Condition (29) depends on how the term $\hat{\mathbf{z}}_t^{\Delta^{(k)}} >> 1$ is large enough to dominate the other term $\frac{c^{(k+1)}}{c^{(k)}}$. The reason for which at a certain step the algorithm stops is because of condition (29): this condition fails at a certain step because the weight of the term $\hat{\mathbf{z}}_t^{\Delta^{(k+1)}}$ decreases at each step k while the weight of the other term $\frac{c^{(k+1)}}{c^{(k)}}$ increases. For the very first steps of the algorithm and for large initial state conditions, the first term is dominant over the second: if $\frac{\partial f_0(v,r)}{\partial r}$ and $\frac{\partial h_0(v,r)}{\partial r}$

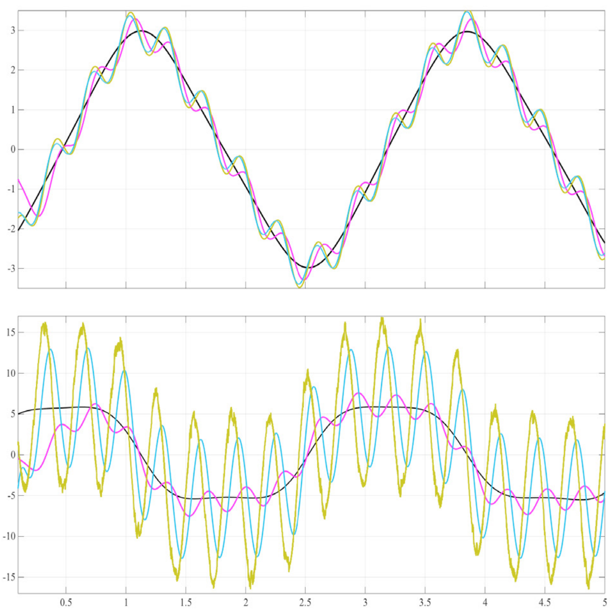

Fig. 1. Sequential estimates of x_1 and x_2

are uniformly bounded then $\Phi(d_\infty, 2c^{(0)} + c) = \Phi(d_\infty, c)$. Moreover, if $\frac{3c^{(0)}+c}{2} \approx c$ then $\Pi^{(1)} \approx \Pi^{(0)}$ condition (29) is largely satisfied at step $k = 0$. In this case from the estimate $\widehat{\mathbf{x}}_t^{(1)}$ we expect a large improvement in the error performances with respect to the estimate $\widehat{\mathbf{x}}_t^{(0)}$ (see simulations in the next section). $\lhd$

6.2 Simulations

Consider the noisy Duffing oscillator

$$\dot{\mathbf{x}}_{1,t} = \mathbf{x}_{2,t},\ \dot{\mathbf{x}}_{2,t} = \mathbf{x}_{1,t} - \mathbf{x}_{1,t}^3 + \mathbf{d}_{1,t},\ \mathbf{y}_t = \mathbf{x}_{1,t} + \mathbf{d}_{2,t} \quad (31)$$

The model disturbance $\mathbf{d}_{1,t}$ is assumed to be constant with values in $[-1, 1]$, while for the measurement disturbance $\mathbf{d}_{2,t}$ we consider the scenario in which it is a sinusoidal disturbance with frequency $\omega \in [10, 20]$ and amplitude $\leqslant 2$. We assumed initial state conditions $x_0 = (-3, 5)^T$.

The state $\mathbf{x}_t$ and its state estimate $\widehat{\mathbf{x}}_t$, obtained from the first K-filter (11), are shown versus time in Fig. 1 (black and light green lines). Notice that since **(A4)** is not guaranteed, the estimate $\widehat{\mathbf{x}}_t$ is very sensitive with respect to measurement noise $\mathbf{d}_{2,t}$ and, at the same time, state magnitude $\|\mathbf{x}_t\|$ (in particular, compare $\mathbf{x}_{2,t}$ and $\widehat{\mathbf{x}}_{2,t}$ in Fig. 1 (black and light green lines).

The second sequential state estimate $\widehat{\mathbf{X}}_t^{(1)}$ is shown versus time in Fig. 1, light blue lines. As you see, $\widehat{\mathbf{X}}_t^{(1)}$ outperforms $\widehat{\mathbf{x}}_t$ by $\approx 45\%$. In particular, notice how the estimate of $\mathbf{x}_{2,t}$ (light blue and light green lines) benefits of a more significant improvement than the estimate of $\mathbf{x}_{1,t}$, (light blue and light green lines). The third sequential state estimate $\widehat{\mathbf{X}}_t^{(2)}$ is shown versus time in Fig. 1 (magenta lines). $\widehat{\mathbf{X}}_t^{(2)}$ outperforms $\widehat{\mathbf{X}}_t^{(1)}$ by $\approx 60\%$.

REFERENCES

V. Andrieu, L. Praly, A. Astolfi, High gain observers with updated gain and homogeneous correction terms, *Automatica*, vol. **81**, pp. 422-428, 2017.

D. Astolfi, R. Postoyan, D. Nesic, Uniting Observers, *IEEE Trans. Autom. Contr.*, DOI 10.1109/TAC.2019.2933395, 2019.

S. Battilotti, Performance optimization via sequential processing for nonlinear state estimation of noisy systems, *IEEE Trans. Autom. Contr.*, submitted, 2019.

S. Battilotti, Robust observer design under measurement noise with gain adaptation and saturated estimates, *Automatica*, vol. **45**, pp. 75-86, 2009.

K. Esfandiari, M. Shakaram, Bank of High-Gain Observers in Output Feedback Control: Robustness Analysis Against Measurement Noise, *IEEE Trans. Syst. Man Cyber.*, vol. **99**, pp. 1-12, 2019.

H. K. Khalil, Cascade high-gain observers in output feedback control, *Automatica*, vol. **80**, pp. 110-118, 2017.

H. K. Khalil and L. Praly, High-gain observers in nonlinear feedback, *Int. Journ. Rob. and Nonlin. Contr.*, vol. **24**, no. **6**, pp. 993-1015, 2014.

M. Kirchman, E.D. Sontag, Y. Wang, Input-output-to-state stability, *SIAM Journ. Contr. and Optim.*, vol. **39**, no. **6**, pp. 1874-1928, 2001.

A. A. Prasov and H. K. Khalil, A Nonlinear High-Gain Observer for Systems With Measurement Noise in a Feedback Control Framework, *IEEE Trans. Autom. Contr.*, vol. **47**, pp. 569-580, 2013.

R. G. Sanfelice, L. Praly, On the performance of high-gain observers with gain adaptation under measurement noise, *Automatica*, vol. **47**, pp. 2165-2176, 2011.

M.M. Seron, J.H. Braslavsky, G.C. Goodwin, *Fundamental Limitations in Filtering and Control*, Springer, 1997.

H. Shim, D. Liberzon, Nonlinear Observers Robust to Measurement Disturbances in an ISS Sense, *IEEE Trans. Autom. Contr.*, vol. **61**, pp. 48-61, 2016.