

Performance optimization via sequential processing for nonlinear state estimation of noisy systems

Stefano Battilotti

Abstract—We propose a framework for designing observers for noisy nonlinear systems with global convergence properties and performing robustness and noise sensitivity. This framework comes out from the combination of a state norm estimator with a chain of filters, adaptively tuned by the state norm estimator. The state estimate is sequentially processed through the chain of filters. Each filter contributes to improve by a certain amount the estimation error performances of the previous filter in terms of noise sensitivity and this amount is quantitatively evaluated using a comparison criterion which considers the ratio of the asymptotic error norm bounds of two consecutive filters in the chain. A recursive algorithm is given for implementing the chain of filters and guaranteeing a sequential error performance optimization process. Simulations show the effectiveness of these chains of filters.

Index Terms—Noisy systems, nonlinear dynamics, observers.

I. INTRODUCTION

DESIGNING state observers with (semi)global convergence properties, enhanced speed performances, robustness and low sensitivity to (measurement) noise is a challenging task. Intrinsic limitations to accomplishing this task have been pointed out for linear systems in [23]. From the past years we have a vast literature on observer design for nonlinear systems (we will not reference this here) focusing on single such performances, almost exclusively convergence domain or speed performances. In many 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 convergence domains (i.e. initial state conditions for which estimation error convergence is guaranteed) without increasing the sensitivity to measurement noise. LMI techniques have been used in [25], [27] and [26] specifically for LPV system. For feedback linearizable systems, large convergence domains and model uncertainties are taken into account in [12], while also additive measurement noise is considered in [19]. In the last contribution, the observer gain switches among two values, one for large state magnitudes (semiglobal observer) and the other for small values (local observer). Also recently in [24] sensitivity of a nonlinear observer versus measurement disturbances has been characterized with the notion of quasi-Disturbance-to-Error Stability (qDES). However, the issue of noise sensitivity reduction is not addressed in the above works. For the same class of nonlinear systems considered in [12] and

[19], very recently [8] and [9] consider a trade-off between convergence speed and the bound on the estimation error, using a bank of HGO observers. For semiglobal observers the effect of measurement disturbances on the ultimate error bound strongly depends on how large the convergence domain is, in other words the observers are designed using the knowledge of the invariant compact set containing the system's trajectories. A consequence is that noise sensitivity can be reduced only at the price of a significant reduction of the convergence domain. Moreover, the observer design depends on the *a priori* given convergence domain: changes in convergence domain require re-designing the observer parameters. In [21] and in [4] the issue of how reducing the sensitivity of a class of observers with global convergence properties to additive measurement noise is discussed for some classes of nonlinear systems with bounded solutions (with unknown bound). However, the error bounds depend on the state magnitude so that for large state initial conditions the state estimate has large excursions and significant deviations from the actual value of the state. More recently, the work [3] unites local observers, which have good error performances versus measurement noise like extended Kalman filters (EKF's) for instance, 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 functioning depends on some local and semiglobal norm estimators together with the exact knowledge of the domains of attraction of the local and semiglobal observers. Moreover, as noticed by the authors, the results may require small disturbances.

In this paper we consider a quite general class of nonlinear systems with model uncertainties (or state noise) and measurement noise and design state observers with the primary objective to optimize the error performances in terms of robustness and noise sensitivity. Important contributions of our work over the existing literature are: 1) global convergence domains (in contrast with *a priori* given and bounded domains as in [19], [8] and [3]), 2) no specific system's structure (in contrast with feedback linearizable systems as in [19], [8] and [9]) taking into account state and measurement noise at the same time and 3) new results on global observers in the absence of noise (in contrast with more restrictive conditions as in [17] and [1]). We do not require boundedness of system's solutions: this has the beneficial effect of obtaining asymptotic error bounds and conditions for improving noise sensitivity not dependent on the convergence domain so that we may improve noise sensitivity without shrinking the estimation error convergence domain. Since good error performances are,

S. Battilotti is with Department of Computer, Control, and Management Engineering "Antonio Ruberti", Sapienza University of Rome, Via Ariosto 25, Italy.

as well known, locally achieved for nonlinear systems by EKF's, we design our observers with a structure the closest as possible to an EKF. In doing this, we follow previous works on Kalman-type filters as for instance [5], [6], [7], [10] and [20], in which globally Lipschitz or linear up to output injection systems (with inputs and no disturbances) are mainly considered. In all these works, the observer gain matrix is updated by the on-line solution of a (differential) Riccati equation. For achieving good results also for systems which are not globally Lipschitz or linear up to output injection the design of the (differential) Riccati equation takes into account also the magnitude of the state through a state norm estimator. The observer we propose in this paper consists of the following parts: I) a state norm estimator (which we call \mathcal{N} -filter) and II) a chain of filters (which we refer to as \mathcal{H} -filters) driven by the \mathcal{N} -filter. Each \mathcal{H} -filter updates its gain matrix using the output of the \mathcal{N} -filter and the solution of a Riccati inequality, similarly to what happens in a steady state Kalman filter. The chain of these filters implements *de facto* a sequential optimization process of the error performances in terms of noise and model uncertainties sensitivity.

II. NOTATION

(I)(vector spaces). \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}_{\geq} (resp. $\mathbb{R}_{>}$) denotes the set of non-negative real numbers (resp. positive real numbers). For any vector $v \in \mathbb{R}^n$ we denote by v_i the i -th element of v . Moreover, $\text{col}(v_1, \dots, v_n)$ is the vector with entries v_1, \dots, v_n and $\text{diag}(v_1, \dots, v_n)$ is the diagonal matrix with diagonal entries v_1, \dots, v_n . \mathbb{C} is the set of complex numbers and $\sigma(S) \subset \mathbb{C}$ denotes the spectrum of $S \in \mathbb{R}^{n \times n}$, with $\lambda_{min}^S := \min_{\lambda \in \sigma(S)} \Re(\lambda)$ and $\lambda_{max}^S := \max_{\lambda \in \sigma(S)} \Re(\lambda)$, where $\Re(\lambda)$ stands for the real part of $\lambda \in \mathbb{C}$. \mathcal{P}_{\geq}^n (resp. $\mathcal{P}_{>}^n$) is the set of symmetric positive definite (resp. semi-definite) matrices $S \in \mathbb{R}^{n \times n}$. Given two symmetric matrices $A, B \in \mathbb{R}^{n \times n}$ we mean by $B \leq A$ that $A - B \in \mathcal{P}_{\geq}^n$.

(II)(weighted and induced norms). $|v|$ denotes the absolute value of $v \in \mathbb{R}$, $\|v\| := \sqrt{v^\top v}$ denotes the euclidean norm of $v \in \mathbb{R}^n$ and the induced norm of $S \in \mathbb{R}^{m \times n}$ is $\|S\| := \sup_{x \in \mathbb{R}^n \setminus \{0\}} (\|Sx\|/\|x\|)$. For each $F \in \mathcal{P}_{>}^m$ and $G \in \mathcal{P}_{>}^n$, the F -weighted norm of $v \in \mathbb{R}^n$ is $\|v\|_F := \sqrt{v^\top F v}$ and the induced (F, G) -weighted norm of $S \in \mathbb{R}^{m \times n}$ is $\|S\|_{(F, G)} := \sup_{x \in \mathbb{R}^n \setminus \{0\}} (\|Sx\|_F/\|x\|_G)$. With some abuse of notation, for $F \in \mathcal{P}_{>}^m$, $G \in \mathcal{P}_{\geq}^n$ and $\gamma > 0$ we write $\|S\|_{(F, G)} \leq \gamma$ by meaning that $\|Sx\|_F \leq \gamma \|x\|_G$ for all $x \in \mathbb{R}^n$ or, equivalently, $S^\top F S \leq \gamma^2 G$.

(III)(monotone functions). Let \mathcal{K}_0 (resp. \mathcal{K}_∞) be the set of continuous non-decreasing (resp. strictly increasing) functions $f : \mathbb{R}_{\geq} \rightarrow \mathbb{R}_{\geq}$ (resp. such that $f(0) = 0$ and $\lim_{s \rightarrow +\infty} f(s) = +\infty$). Let \mathcal{K}_∞^1 be the set of functions $f \in \mathcal{K}_\infty$ continuously differentiable on $\mathbb{R}_{>}$ and such that $0 < \inf_{s>0} \frac{s}{f(s)} \frac{df}{ds}(s)$ (for instance \mathcal{K}_∞ algebraic functions are all in \mathcal{K}_∞^1). Let \mathcal{L}_∞ be the set of continuous strictly decreasing functions $f : \mathbb{R}_{>} \rightarrow \mathbb{R}_{>}$ such that $\lim_{s \rightarrow +\infty} f(s) = 0$. Let \mathcal{L}_∞^1 be the set of functions $f \in \mathcal{L}_\infty$ continuously differentiable on $\mathbb{R}_{>}$ and such that

$\sup_{s>0} \frac{s}{f(s)} \frac{df}{ds}(s) < 0$ (for instance \mathcal{L}_∞ algebraic functions are all in \mathcal{L}_∞^1).

(IV)(saturation functions). A saturation function $\text{sat}_c : \mathbb{R}^n \rightarrow \mathbb{R}^n$ with saturation level $c > 0$ is a function $\text{sat}_c(x) := (\text{sat}_c(x_1), \dots, \text{sat}_c(x_n))^\top$, $x \in \mathbb{R}^n$, such that for each $i = 1, \dots, n$ and $x_i \in \mathbb{R}$:

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

Although this function is not continuously differentiable at all points, it is possible to smooth it out. In this paper we will consider at least once continuously differentiable such saturation functions with no further comment.

III. CLASS OF SYSTEMS

Consider a nonlinear system of the general form

$$\begin{aligned} \dot{x} &= f(x, d) := Ax + \mathcal{F}(x, d), \\ y &= h(x, d) := Cx + \mathcal{H}(x, d) \end{aligned} \quad (2)$$

with state $x \in \mathbb{R}^n$, output $y \in \mathbb{R}^p$, \mathcal{F} and \mathcal{H} continuously differentiable functions with $\frac{\partial \mathcal{F}}{\partial x}(0, 0) = 0$ and $\frac{\partial \mathcal{H}}{\partial x}(0, 0) = 0$, and disturbances (or exogenous inputs) $d \in \mathcal{D}$, the space of (piecewise) continuous bounded functions $d : \mathbb{R}_{\geq} \rightarrow D \subset \mathbb{R}^s$ with sup norm $\|d\|_\infty := \sup_{t \in \mathbb{R}_{\geq}} \|d(t)\|$ uniformly bounded by a known $d_\infty > 0$. We will denote $x(t), y(t)$ and $d(t)$ by $\mathbf{x}_t, \mathbf{y}_t$ and, respectively, \mathbf{d}_t , with initial state $x(0) = x_0$ at $t = 0$. More precisely, $\mathbf{x}_t(x, s; \mathbf{d})$ (resp. $\mathbf{y}_t(x, s; \mathbf{d})$) will denote the value at time t of the unique solution (resp. output) of system (2) with input \mathbf{d} and initialized at x at time s , i.e. $\mathbf{x}_s(x, s; \mathbf{d}) = x$. Our problem is to design state estimators for (2) with good performances in terms of noise sensitivity. Throughout the paper, we assume forward completeness of (2).

(FWC). (Forward completeness) The solutions \mathbf{x}_t of (2) are defined for all $(x_0, \mathbf{d}, t) \in \mathbb{R}^n \times \mathcal{D} \times \mathbb{R}_{\geq}$.

Completeness is a severe restriction. However, as well-known, a necessary condition for the existence of an observer for (2) is the forward unboundedness observability property (FUO). Any system (2) satisfying the (FUO) property also satisfies (FWC) after a suitable change of the time scale.

IV. AN OVERVIEW OF THE ESTIMATION FRAMEWORK

In this introductory section we describe the estimation framework and how it works, leaving the technicalities to be discussed in the next sections. The first element in our framework is a state norm estimator (which, in what follows, we refer to as \mathcal{N} -filter), in the sense that its output is some variable $\hat{\mathbf{v}}_t$ from which an upper bound of the state norm $\|\mathbf{x}_t\|$ is determined as follows: for some $\beta \in \mathcal{K}_\infty^1$ and $\hat{\gamma} > 1$ we have

$$\|\mathbf{x}_t\| \leq \beta^{-1}(\hat{\mathbf{v}}_t + \hat{\gamma}) \quad (3)$$

for all $(x_0, \mathbf{d}, t) \in \mathbb{R}^n \times \mathcal{D} \times [t_0, +\infty)$ (t_0 depending on the initial state x_0). The \mathcal{N} -filter has the following parametric structure:

$$\dot{\hat{\mathbf{v}}}_t = -\lambda \hat{\mathbf{v}}_t + \alpha(\|\mathbf{y}_t\|) + \delta(d_\infty), \quad \hat{\mathbf{v}}_0 > 0, \quad (4)$$

with $\lambda \in (0, 1)$. The functions $\delta, \alpha \in \mathcal{K}_0$ and $\lambda \in (0, 1)$ are obtained from a differential inequality of the form

$$\dot{v}(\mathbf{x}_t, t) \leq -\lambda v(\mathbf{x}_t, t) + \alpha(\|\mathbf{y}_t\|) + \delta(\|\mathbf{d}_t\|) \quad (5)$$

which must hold along the solutions of (2) for some continuously differentiable function $v : \mathbb{R}^n \times \mathbb{R}_{\geq} \rightarrow \mathbb{R}$ and for all $(x_0, \mathbf{d}, t) \in \mathbb{R}^n \times \mathcal{D} \times \mathbb{R}_{\geq}$. On the other hand, the function $\beta \in \mathcal{K}_{\infty}^1$, which is instrumental in (3) to determine an upper bound for the state norm $\|\mathbf{x}_t\|$, is related to $v(x, t)$ as follows: there exists a time $\bar{t} > 0$ for which $v(x, t) \geq \beta(\|x\|)$ for all $(x, t) \in \mathbb{R}^n \times [\bar{t}, +\infty)$. In section V we will discuss existence conditions (Proposition 5.1) and constructive methodologies (Propositions 5.2 and 5.3) for the functions v, β, δ and α with the above mentioned properties.

The second element in our framework is a set of filters (which, in what follows, we refer to as \mathcal{K} -filters) adaptively tuned by the \mathcal{N} -filter and sequentially connected to each other so that to form a dynamical chain. Each $\mathcal{K}^{(k)}$ -filter, $k \geq 1$, of the chain is specifically designed for estimating the state of (2) as follows. The variable $\hat{\mathbf{X}}_t^{(k-1)}$ (we are setting $\hat{\mathbf{X}}_t^{(0)} := 0$) keeps track of the state estimate of (2) up to the $\mathcal{K}^{(k-1)}$ -filter. The $\mathcal{K}^{(k-1)}$ -filter feeds forward $\hat{\mathbf{X}}_t^{(k-1)}$ and $\hat{\mathbf{X}}_t^{(k-1)}$ to the $\mathcal{K}^{(k)}$ -filter (the next filter in the dynamical chain), which estimates the error $\mathbf{e}_t^{(k)} := \mathbf{x}_t - \hat{\mathbf{X}}_t^{(k-1)}$ described by the equations:

$$\begin{aligned} \dot{\mathbf{e}}_t^{(k)} &= A\mathbf{e}_t^{(k)} + \hat{\Psi}_t^{(k)} + \mathcal{F}(\mathbf{e}_t^{(k)} + \hat{\mathbf{X}}_t^{(k-1)}, \mathbf{d}_t), \\ \mathbf{w}_t^{(k)} &= C\mathbf{e}_t^{(k)} + \mathcal{H}(\mathbf{e}_t^{(k)} + \hat{\mathbf{X}}_t^{(k-1)}, \mathbf{d}_t), \end{aligned} \quad (6)$$

with $\hat{\Psi}_t^{(k)} := A\hat{\mathbf{X}}_t^{(k-1)} - \dot{\hat{\mathbf{X}}}_t^{(k-1)}$ and output error $\mathbf{w}_t^{(k)} := \mathbf{y}_t - C\hat{\mathbf{X}}_t^{(k-1)}$. Consistently with the fact that $\hat{\mathbf{X}}_t^{(0)} := 0$, the first filter of the chain, i.e. the $\mathcal{K}^{(1)}$ -filter, is designed on (2) and computes a first estimate $\hat{\mathbf{X}}_t^{(1)}$ of \mathbf{x}_t .

The $\mathcal{K}^{(k)}$ -filter estimates $\mathbf{e}_t^{(k)}$ with the variable $\hat{\mathbf{e}}_t^{(k)}$ and updates the state estimate of (2) as $\hat{\mathbf{X}}_t^{(k)} := \hat{\mathbf{e}}_t^{(k)} + \hat{\mathbf{X}}_t^{(k-1)}$. For the $\mathcal{K}^{(k)}$ -filter we propose the following parametric structure:

$$\begin{aligned} \dot{\hat{\mathbf{e}}}_t^{(k)} &= A\hat{\mathbf{e}}_t^{(k)} + \hat{\Psi}_t^{(k)} + \widehat{\mathcal{F}}^{(k)}(\hat{\mathbf{e}}_t^{(k)}, \hat{\mathbf{X}}_t^{(k-1)}, \Lambda^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})) \\ &\quad + \mathcal{H}^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})(\mathbf{w}_t^{(k)} - \hat{\mathbf{w}}_t^{(k)}), \\ \hat{\mathbf{w}}_t^{(k)} &= C\hat{\mathbf{e}}_t^{(k)} + \widehat{\mathcal{H}}^{(k)}(\hat{\mathbf{e}}_t^{(k)}, \hat{\mathbf{X}}_t^{(k-1)}, \Lambda^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})), \end{aligned} \quad (7)$$

with $\hat{\mathbf{v}}_t$ the output of the \mathcal{N} -filter, $\hat{\gamma} > 1$ and

$$\mathcal{K}^{(k)}(z) := \frac{\gamma_A^{(k)}(z)}{(\gamma_C^{(k)}(z))^2} \mathcal{P}^{(k)-1}(z) C^\top R^{-1} (e_{d_\infty}^{(k)} + \hat{X}_{d_\infty}^{(k)}), \quad (8)$$

$$\mathcal{P}^{(k)}(z) := \Lambda^\top(z) \Pi^{(k)} \Lambda(z),$$

$$\widehat{\mathcal{F}}^{(k)}(\hat{e}, \hat{X}, \Lambda) := \mathcal{F} \left(\Lambda^{-1} \text{sat}_{e_{d_\infty}^{(k)}}(\Lambda \hat{e}) + \Lambda^{-1} \text{sat}_{\hat{X}_{d_\infty}^{(k)}}(\Lambda \hat{X}), 0 \right),$$

$$\widehat{\mathcal{H}}^{(k)}(\hat{e}, \hat{X}, \Lambda) := \mathcal{H} \left(\Lambda^{-1} \text{sat}_{e_{d_\infty}^{(k)}}(\Lambda \hat{e}) + \Lambda^{-1} \text{sat}_{\hat{X}_{d_\infty}^{(k)}}(\Lambda \hat{X}), 0 \right),$$

where the admissible choices for the design parameters $e_{d_\infty}^{(k)}, \hat{X}_{d_\infty}^{(k)} > 0$ (depending only on d_∞), $\Pi^{(k)} \in \mathcal{P}_{>}^n$ and functions $R : \mathbb{R}_{>} \rightarrow \mathcal{P}_{>}^p$, $\Lambda^{(k)} : \mathbb{R}_{>} \rightarrow \mathbb{R}^{n \times n}$, $\gamma_A^{(k)} \in \mathcal{K}_0$ and $\gamma_C^{(k)} \in \mathcal{K}_\infty$ are discussed in section VI-B relying on a certain number of assumptions on A, C, \mathcal{F} and \mathcal{H} . Estimation

error convergence is discussed in Theorem 6.1 (for a $\mathcal{K}^{(k)}$ -filter designed on (6) with $k \geq 2$) and, respectively, in Theorem 6.2 (for the $\mathcal{K}^{(1)}$ -filter, which is designed on (2)). A first result ((58), Theorem 6.1) states that, under the given assumptions on A, C, \mathcal{F} and \mathcal{H} , the norm of the estimation error $\mathbf{e}_t^{(k+1)} := \mathbf{x}_t - \hat{\mathbf{X}}_t^{(k)}$ satisfies an asymptotic bound of the form:

$$\limsup_{t \rightarrow +\infty} \frac{\|\Lambda^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})\mathbf{e}_t^{(k+1)}\|}{\gamma_d^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})} < c_{d_\infty}^{(k+1)}, \quad (9)$$

where $c_{d_\infty}^{(k+1)} > 0$ (depending only on d_∞) and $\gamma_d^{(k)} \in \mathcal{L}_\infty^1$ are directly obtained from the given assumptions. From (9) it also follows that

$$\limsup_{t \rightarrow +\infty} \frac{\|\mathbf{e}_t^{(k+1)}\|}{\omega^{(k)}(\hat{\mathbf{v}}_t)} < c_{d_\infty}^{(k+1)} \quad (10)$$

where $\omega^{(k)} \in \mathcal{K}_0$ is defined as

$$\omega^{(k)}(s) := \gamma_d^{(k)}(s + \hat{\gamma}) / \sqrt{\lambda_{\min}^{\Lambda^{(k)\top}(s + \hat{\gamma})\Lambda^{(k)}(s + \hat{\gamma})}}.$$

The bound (10) provides a rough measure of how poor the asymptotic performances of the state estimate $\hat{\mathbf{X}}_t^{(k)}$ are in terms of noise sensitivity and sensitivity to initial state conditions. A second important result ((59), Theorem 6.1) states that if, in addition,

$$\sup_{s \geq 1} \omega^{(k)}(s) := \gamma_\infty^{(k)} < +\infty \quad (11)$$

then

$$\limsup_{t \rightarrow +\infty} \|\mathbf{e}_t^{(k+1)}\| < \gamma_\infty^{(k)} c_{d_\infty}^{(k+1)}. \quad (12)$$

The bound (12) corresponds to a more desirable situation in terms of noise sensitivity, indeed the bound on the estimation error norm is not sensitive to initial state conditions and depends only on d_∞ . Also, this bound may be optimized by properly tuning the filter's design parameters. However, in contrast with our assumptions on A, C, \mathcal{F} and \mathcal{H} , condition (11) is somewhat restrictive: it holds, for instance, when \mathcal{F} and \mathcal{H} are globally Lipschitz but also in many other cases (example 6.1). We observed that there are specific structures of \mathcal{F} and \mathcal{H} in which (11) may fail (see a detailed discussion in remark 6.9 with slight modifications of example 6.1). In view of this, the main purpose of the chain of \mathcal{K} -filters is to improve sequentially the error performances of the $\mathcal{K}^{(1)}$ -filter when (11) fails for $k = 1$. By improving the error performances we mean guaranteeing tighter asymptotic bounds (10) on the estimation error norm. For sequentially optimizing the error performances of the $\mathcal{K}^{(1)}$ -filter we proceed as follows: first, for each $\mathcal{K}^{(k)}$ -filter we design $\Lambda^{(k)}$ so as to slow down the growth rate of $\omega^{(k)}$ and tend (as the number k of \mathcal{K} -filters increases) to the ‘‘limit’’ condition (11). At the same time, for guaranteeing an actual improvement of the asymptotic bound (10) (as k increases), we compare the error performances of $\hat{\mathbf{X}}_t^{(k)}$ and, respectively, $\hat{\mathbf{X}}_t^{(k-1)}$ as follows:

$$\limsup_{t \rightarrow +\infty} \frac{\omega^{(k)}(\hat{\mathbf{v}}_t) c_{d_\infty}^{(k+1)}}{\omega^{(k-1)}(\hat{\mathbf{v}}_t) c_{d_\infty}^{(k)}} = r^{(k)} \quad (13)$$

i.e. by comparing the corresponding asymptotic error bounds (10). As a consequence of (13), the state estimation error bound of $\hat{\mathbf{X}}_t^{(k)}$ is asymptotically reduced by a factor $r^{(k)}$ if compared with that of $\hat{\mathbf{X}}_t^{(k-1)}$ and if $r^{(k)} < 1$ we say that $\hat{\mathbf{X}}_t^{(k)}$ outperforms $\hat{\mathbf{X}}_t^{(k-1)}$ by $(1 - r^{(k)}) \times 100(\%)$ (section VII-A, definition 7.1). In this sense we mean that the noise sensitivity of $\hat{\mathbf{X}}_t^{(k)}$ is improving by $(1 - r^{(k)}) \times 100(\%)$ in comparison with $\hat{\mathbf{X}}_t^{(k-1)}$. In our framework, the condition $r^{(k)} < 1$ is achieved by properly designing $e_{d_\infty}^{(k)}$, $\hat{X}_{d_\infty}^{(k)}$ and $\Pi^{(k)}$ in (8).

In section VII-C, under the same assumptions on (2) as of Theorem 6.2 and recalled at the beginning of section VII-B for sake of clarity, we give the details of a recursive algorithm for designing the chain of \mathcal{H} -filters in such a way that each \mathcal{H} -filter outperforms its predecessor in the chain. The algorithm starts at step I with the design of the \mathcal{N} -filter and the $\mathcal{H}^{(1)}$ -filter on the system (2) (according to Theorem 6.2) with the initial state estimate $\hat{\mathbf{X}}_t^{(1)}$; in the following recursive steps II-III a $\mathcal{H}^{(k)}$ -filter ($k \geq 2$) is designed on the system (6) (according to Theorem 6.1) using $\hat{\mathbf{X}}_t^{(k-1)}$ and updating the state estimate as $\hat{\mathbf{X}}_t^{(k)}$. Assumptions of Theorem 6.1 on (6) are guaranteed by design of the previous $\mathcal{H}^{(k-1)}$ -filter in the chain and the initial assumptions on (2). Finally, at step IV structural conditions on the design parameters of the $\mathcal{H}^{(k)}$ -filter are given for $r^{(k)} < 1$. The algorithm is iterated for the next $\mathcal{H}^{(k+1)}$ -filter by jumping back to step II.

V. DESIGN OF STATE NORM ESTIMATORS

The first issue we discuss in detail is the design of a state norm estimator (i.e. a \mathcal{N} -filter) for (2). To this aim, we formulate the following assumption:

(SNE). (State norm estimators) *There exist a continuously differentiable function $v : \mathbb{R}^n \times \mathbb{R}_{\geq} \rightarrow \mathbb{R}$, $\lambda > 0$ and $\delta, \alpha \in \mathcal{K}_0$ such that*

$$\begin{aligned} \text{(PDI)} : \quad & \frac{\partial v}{\partial x}(x, t) f(x, \mathbf{d}_t) + \frac{\partial v}{\partial t}(x, t) \\ & \leq -\lambda v(x, t) + \alpha(\|h(x, \mathbf{d}_t)\|) + \delta(\|\mathbf{d}_t\|), \end{aligned} \quad (14)$$

for all $(x, \mathbf{d}, t) \in \mathbb{R}^n \times \mathcal{D} \times \mathbb{R}_{\geq}$. In addition, there exist $\beta \in \mathcal{K}_\infty^1$ and $\bar{t} \geq 0$ such that for all $(x, t) \in \mathbb{R}^n \times [\bar{t}, +\infty)$

$$\text{(ULB)} : v(x, t) \geq \beta(\|x\|). \quad (15)$$

We also say that the tuple $(v, \lambda, \alpha, \delta, \beta, \bar{t})$ satisfies a (SNE) condition and we use separate terminologies as well: $(v, \lambda, \alpha, \delta)$ satisfies a (PDI) condition or (v, β, \bar{t}) satisfies a (ULB) condition. Moreover, we directly assume $\lambda \in (0, 1]$ (otherwise, we consider $\min\{\lambda, 1\}$ in place of λ in (PDI)).

The function $v(x, t)$ satisfying (SNE) is reminiscent of exponential IOSS-Lyapunov function previously introduced in [14]: in contrast, our function $v(x, t)$ is time-varying and lower bounded by a \mathcal{K}_∞^1 -class function of $\|x\|$ after some finite time \bar{t} . As in [14], the interest in (14), (15) is motivated by the following result.

Proposition 5.1: *Assume $(v, \lambda, \alpha, \delta, \beta, \bar{t})$ satisfies a (SNE) condition. For each $(x_0, \mathbf{d}) \in \mathbb{R}^n \times \mathcal{D}$ and $\hat{\gamma} > 0$ there exists $t_0 \geq \bar{t}$ such that for $t \geq t_0$:*

$$\|\mathbf{x}_t(x_0, 0; \mathbf{d})\| \leq \beta^{-1}(\hat{\mathbf{v}}_t(\hat{v}_0, 0; \mathbf{y}) + \hat{\gamma}), \quad (16)$$

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

$$\begin{aligned} \dot{\hat{\mathbf{v}}}_t(\hat{v}_0, 0; \mathbf{y}) &= -\lambda \hat{\mathbf{v}}_t(\hat{v}_0, 0; \mathbf{y}) + \alpha(\|\mathbf{y}_t(x_0, 0; \mathbf{d})\|) \\ &+ \delta(d_\infty), \quad \hat{v}_0 > 0. \end{aligned} \quad (17)$$

Proof. On account of (14) for all $(x_0, \mathbf{d}, t) \in \mathbb{R}^n \times \mathcal{D} \times \mathbb{R}_{\geq}$

$$\begin{aligned} \frac{d}{dt} v(\mathbf{x}_t(x_0, 0; \mathbf{d}), t) &\leq -\lambda v(\mathbf{x}_t(x_0, 0; \mathbf{d}), t) \\ &+ \alpha(\|h(\mathbf{x}_t(x_0, 0; \mathbf{d}), \mathbf{d}_t)\|) + \delta(\|\mathbf{d}_t\|). \end{aligned}$$

By subtracting (17) from the latter and using Gronwall's inequality, we get for all $(x_0, \mathbf{d}, t) \in \mathbb{R}^n \times \mathcal{D} \times \mathbb{R}_{\geq}$

$$v(\mathbf{x}_t(x_0, 0; \mathbf{d}), t) \leq \hat{\mathbf{v}}_t(\hat{v}_0, 0; \mathbf{y}) + e^{-\lambda t} |v(x_0, 0) - \hat{v}_0|. \quad (18)$$

Inequality (16) follows directly from (15). \triangleleft

Proposition 5.1 proves that a (SNE) condition guarantees the existence of a \mathcal{N} -filter, defined in (17), and precisely states in which sense, specified by (16), a state norm estimate has to be meant. In this sense we say that a tuple $(v, \lambda, \alpha, \delta, \beta, \bar{t})$ satisfies a (SNE) condition with associated \mathcal{N} -filter (17) and state norm estimate (16).

We list below some useful properties:

Property (P1). *Let $(v, \lambda, \alpha, \delta, \beta, \bar{t})$ satisfies a (SNE) condition. Any tuple $(v + b, \lambda, \alpha, \delta + \lambda b, \beta, \bar{t})$, with $b > 0$, still satisfies a (SNE) condition (this trivially follows from the definition).*

Property (P2). *Let $(v, \lambda, \alpha, \delta)$ satisfies a (PDI) condition and $v(x, t) \geq \beta(\|x\|) - b$ for some $b, \bar{t} > 0$, $\beta \in \mathcal{K}_\infty^1$ and for all $(x, t) \in \mathbb{R}^n \times [\bar{t}, +\infty)$. The tuple $(v + b, \lambda, \alpha, \delta + \lambda b, \beta, \bar{t})$ satisfies a (SNE) condition (this follows from Property (P1)).*

Property (P3). *Given $\mu, \nu > 0$, it is possible to re-design a tuple $(v, \lambda, \alpha, \delta, \beta, \bar{t})$ satisfying a (SNE) condition into a new tuple $(\tilde{v}, \tilde{\lambda}, \tilde{\alpha}, \tilde{\delta}, \tilde{\beta}, \tilde{t})$, with $\beta(s) := \nu s^\mu$, still satisfying a (SNE) condition (the proof is found in section A of the appendix).*

Remark 5.1: The presence of $\bar{t} \geq 0$ in a tuple $(v, \lambda, \alpha, \delta, \beta, \bar{t})$ satisfying a (SNE) condition is a consequence of the fact that $v(x, t)$ is time-varying. As it will be shown in section V-B (see assumption (BWR)), \bar{t} may be very small but nonzero and it can be interpreted as a time lapse needed for reconstructing the initial state condition x from (backward) output trajectories. For having $\bar{t} = 0$ we have to look for stationary (i.e. time-invariant) solutions $v(x)$ of the (PDI) condition (under more restrictive conditions). \triangleleft

Example 5.1: (Linear systems). Consider a linear system

$$\dot{x} = f(x, d) := Ax + Bd, \quad y = h(x, d) := Cx + Pd$$

with detectable (C, A) . Let $L \in \mathbb{R}^{n \times p}$ be such that $\sigma(A - LC) \subset \mathbb{C}^-$ and $Q, \Pi \in \mathcal{P}_>^n$ be such that $\Pi(A - LC) + (A - LC)^\top \Pi = -Q$. If $v(x) := x^\top \Pi x$ we have for any $a_0, d_0 \in \mathbb{R}_{>}$ and for all $(x, \mathbf{d}, t) \in \mathbb{R}^n \times \mathcal{D} \times \mathbb{R}_{\geq}$

$$\begin{aligned} \frac{\partial v}{\partial x}(x) f(x, \mathbf{d}_t) &\leq -\frac{\lambda_{\min}^Q}{\lambda_{\max}^{\Pi}} v(x) + \alpha(\|h(x, \mathbf{d}_t)\|) + \delta(\|\mathbf{d}_t\|) \\ &+ x^\top \Pi \left(\frac{1}{a_0} LL^\top + \frac{1}{d_0} (B - LP)(B - LP)^\top \right) \Pi x \end{aligned} \quad (19)$$

with $\alpha(s) := a_0 s^2$ and $\delta(s) := d_0 s^2$. If, in addition, $a_0, d_0 > 0$ are chosen so that $\lambda_0 := \frac{\lambda_{min}^Q}{\lambda_{max}^S} - \lambda_{max}^S > 0$ where

$$S := \Pi^{1/2} \left(\frac{1}{a_0} LL^\top + \frac{1}{d_0} (B - LP)(B - LP)^\top \right) \Pi^{1/2}, \quad (20)$$

then $(v, \lambda_0, \alpha, \delta, \beta, 0)$ satisfy a (SNE) condition with $\beta(s) := \lambda_{min}^\Pi s^2$. \triangleleft

Example 5.2: Consider the (Van Der Pol) system

$$\begin{aligned} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} &= f(x, d) := \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 0 \\ x_1^2 x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ x_1 \end{pmatrix} d_1, \\ y &= h(x, d) := x_1 + d_2. \end{aligned} \quad (21)$$

Let $Q, \Pi \in \mathbb{P}_{>}^2$ be such that

$$\Pi \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}^\top \Pi = -Q \quad (22)$$

and $a_1, a_2 > 0$ be such that $\lambda_0 := \frac{\lambda_{min}^Q}{\lambda_{max}^S} - \lambda_{max}^S > 0$ where

$$S := \Pi^{1/2} \left(\frac{1}{a_1} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \frac{1}{a_2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right) \Pi^{1/2}.$$

Define $v(x) := z^\top(x) \Pi z(x)$, where $z(x) := (x_1, x_2 - 2x_1 + (1/3)x_1^3)^\top$. There exist $b, \beta_0 > 0$ such that $v(x) \geq \beta_0 \|x\|^{2/3} - b$ for all $x \in \mathbb{R}^2$. It follows from Property (P2) that $(v + b, \lambda_0, \alpha, \delta + b\lambda_0, \beta, 0)$ satisfy a (SNE) condition with $\beta(s) := \beta_0 s^{2/3}$ and 6-th degree polynomials $\alpha, \delta \in \mathcal{K}_0$ (we leave the computations to the reader). \triangleleft

A. Closed-form solutions of (PDI)

For obtaining closed-form solutions of (PDI), we discard finite exit times from \mathbb{R}^n for the backward solutions of (2).

(BWC). (Backward completeness) The solutions $\mathbf{x}_s(x, t; \mathbf{d})$ of (2) are defined for all $(x, \mathbf{d}, t, s) \in \mathbb{R}^n \times \mathcal{D} \times \mathbb{R}_{\geq} \times [0, t]$.

In what follows we will use the notation

$$\mathbf{Y}_s(x, t; \mathbf{d}) := h(\mathbf{x}_s(x, t; \mathbf{d}), \mathbf{d}_s) - h(0, \mathbf{d}_s). \quad (23)$$

Proposition 5.2: Under assumption (BWC) and for any given $\lambda > 0$ and $\alpha \in \mathcal{K}_0$, $(v, \lambda, \bar{\alpha}, \bar{\delta})$ satisfies a (PDI) condition with

$$v(x, t) := \int_0^t e^{-\lambda(t-s)} \alpha(\|\mathbf{Y}_s(x, t; \mathbf{d})\|) ds \quad (24)$$

and

$$\bar{\alpha}(s) := \alpha(2s), \quad \bar{\delta}(s) := \sup_{\|d\| \leq s} \alpha(2\|h(0, d)\|). \quad (25)$$

Proof. Notice that for any $\Delta t \in \mathbb{R}$

$$\mathbf{x}_s(\mathbf{x}_{t+\Delta t}(x, t; \mathbf{d}), t + \Delta t; \mathbf{d}) = \mathbf{x}_s(x, t; \mathbf{d}) \quad (26)$$

We have

$$\begin{aligned} \frac{v(\mathbf{x}_{t+\Delta t}(x, t; \mathbf{d}), t + \Delta t) - v(x, t)}{\Delta t} &= \left(\frac{e^{-\lambda \Delta t} - 1}{\Delta t} \right) v(x, t) \\ &+ \frac{e^{-\lambda \Delta t}}{\Delta t} \int_t^{t+\Delta t} e^{-\lambda(t-s)} \alpha(\|\mathbf{Y}_s(x, t; \mathbf{d})\|) ds. \end{aligned}$$

Letting Δt tend to 0 we get

$$\begin{aligned} &\frac{\partial v}{\partial x}(x, t) f(x, \mathbf{d}_t) + \frac{\partial v}{\partial t}(x, t) \\ &= -\lambda v(x, t) + \alpha(\|h(x, \mathbf{d}_t) - h(0, \mathbf{d}_t)\|) \\ &\leq -\lambda v(x, t) + \bar{\alpha}(\|h(x, \mathbf{d}_t)\|) + \bar{\delta}(\|\mathbf{d}_t\|), \end{aligned} \quad (27)$$

where we used $\alpha(s+r) \leq \alpha(2s) + \alpha(2r)$ for all $s, r \geq 0$. \triangleleft

B. Sufficient conditions for (ULB)

In order to have the function $v(x, t)$ in (24) uniformly lower bounded as specified in the (ULB) condition, we invoke a kind of uniform (backward) state reconstructibility property from the outputs.

(BWR). (Uniform backward reconstructibility). There exist $\bar{t}, b > 0$, $\alpha \in \mathcal{K}_0$ and $\beta \in \mathcal{K}_\infty^1$ such that for all $(x, \mathbf{d}) \in \mathbb{R}^n \times \mathcal{D}$

$$\int_0^{\bar{t}} \alpha(\|\mathbf{Y}_s(x, \bar{t}; \mathbf{d})\|) ds \geq \beta(\|x\|) - b. \quad (28)$$

A similar reconstructibility property was used in [22] in a noise-free context for state-dependent solutions of differential Riccati equations. For linear systems (2): $\dot{x} = Ax + \mathcal{F}d$, $y = Cx + \mathcal{H}d$, assumption (BWR) is strictly related to observability under unknown inputs d ([16]).

Proposition 5.3: Under assumptions (BWC) and (BWR) and for any given $\lambda > 0$ there exist $\bar{t}, \bar{b} > 0$ and $\bar{\beta} \in \mathcal{K}_\infty^1$ such that $(v + \bar{b}, \lambda, \bar{\alpha}, \bar{\delta} + \lambda \bar{b}, \bar{\beta}, \bar{t})$ satisfies a (SNE) condition with $v : \mathbb{R}^n \times \mathbb{R}_{\geq}$ defined in (24) and $\bar{\alpha}, \bar{\delta} \in \mathcal{K}_0$ defined in (25).

Proof. Let $\bar{t}, b > 0$, $\alpha \in \mathcal{K}_0$ and $\beta \in \mathcal{K}_\infty^1$ be as in assumption (BWR) and $v : \mathbb{R}^n \times \mathbb{R}_{\geq}$ as in (24). For all $t \geq \bar{t}$ we have

$$\begin{aligned} v(x, t) &\geq \int_{t-\bar{t}}^t e^{-\lambda(t-s)} \alpha(\|\mathbf{Y}_s(x, t; \mathbf{d})\|) ds \\ &\geq e^{-\lambda \bar{t}} \int_0^{\bar{t}} \alpha(\|\mathbf{Y}_\vartheta(x, \bar{t}; \bar{\mathbf{d}})\|) d\vartheta \end{aligned} \quad (29)$$

and $\bar{\mathbf{d}}_\vartheta := \mathbf{d}_{\vartheta+t-\bar{t}}$. But $\bar{\mathbf{d}} \in \mathcal{D}$ and on account of assumption (BWR)

$$v(x, t) \geq e^{-\lambda \bar{t}} (\beta(\|x\|) - b) := \bar{\beta}(\|x\|) - \bar{b} \quad (30)$$

for all $(x, t) \in \mathbb{R}^n \times [\bar{t}, +\infty)$, which proves that $(v + \bar{b}, \bar{\beta}, \bar{t})$ satisfy a (ULB) condition. On the other hand, $(v + \bar{b}, \lambda, \bar{\alpha}, \bar{\delta} + \lambda \bar{b})$ satisfies a (PDI) condition by proposition 5.2 and Property (P2). \triangleleft

From (BWR) we see that \bar{t} may be very small but nonzero. More conditions for (BWR) based on Lie derivatives will be given in other related papers.

C. A stronger (SNE) condition

We have seen how the (SNE) condition has a key role in the design of a \mathcal{N} -filter. However, when the \mathcal{N} -filter is coupled with another filter (like an observer) it is better to have $\alpha(\|h(x, d)\|)$ not increasing faster than $v(x, t)$. For this reason, we introduce a slightly stronger (SNE) condition:

(SSNE). There exists a tuple $(v, \lambda, \alpha, \delta, \beta, \bar{t})$ satisfying a (SNE) condition and, in addition, there exist $\zeta > 0$ and $\xi \in \mathcal{K}_0$ such that for all $(x, d) \in \mathbb{R}^n \times \mathcal{D}$ and $t \geq \bar{t}$

$$(\mathbf{UUB}) : \alpha(\|h(x, d)\|) \leq \zeta v(x, t) + \xi(\|d\|). \quad (31)$$

Using a similar terminology to the previous sections, we say that a tuple $(v, \lambda, \alpha, \delta, \beta, \zeta, \xi, \bar{t})$ satisfies a (SSNE) condition (with associated \mathcal{N} -filter (17)) or, referring only to (31), that $(v, \alpha, \zeta, \xi, \bar{t})$ satisfies a (UUB) condition. Notice that, on account of the (ULB) condition, there always exist $\zeta \in \mathcal{K}$ and $\xi \in \mathcal{K}_0$ such that for all $(x, d) \in \mathbb{R}^n \times D$ and $t \geq \bar{t}$

$$\alpha(\|h(x, d)\|) \leq \zeta(v(x, t)) + \xi(\|d\|). \quad (32)$$

The (UUB) condition requires a *linear*¹ $\zeta \in \mathcal{K}_0$.

Example 5.1 (cont'ed). The tuple $(v, \lambda, \alpha, \delta, \beta, \zeta, \xi, 0)$, where $(v, \lambda, \alpha, \delta, \beta, 0)$ is the tuple which satisfy a (SNE) condition in example 5.1, satisfy a (SSNE) condition with $\zeta := \frac{2a_0\|C\|^2}{\lambda_{\min}^{\Pi}}$ and $\xi(s) := 2a_0\|Q\|^2s^2$, with $a_0 > 0$ given in (20). \triangleleft

Example 5.2 (cont'ed). The tuple $(v, \lambda, \alpha, \delta, \beta, \zeta, \xi, 0)$, where $(v, \lambda, \alpha, \delta, \beta, 0)$ is the tuple which satisfy a (SNE) condition in example 5.2, cannot satisfy a (SSNE) condition for any $\zeta > 0$. We modify $v(x)$ in example 5.2 as follows:

$$\tilde{v}(x) := v(x) + \frac{1}{2} \begin{pmatrix} 0 & 1 \end{pmatrix} \Pi \begin{pmatrix} 1 \\ 2 \end{pmatrix} x_1^4 \quad (33)$$

with $z(x) := (x_1, x_2 - 2x_1)^\top$ and Π given in (22). We obtain a tuple $(\tilde{v}, \lambda, \tilde{\alpha}, \delta, \beta, \zeta, \xi, 0)$ satisfying a (SSNE) condition, with 4-th degree polynomials $\tilde{\alpha}, \tilde{\delta}, \xi \in \mathcal{K}_0$. \triangleleft

The following property holds for a tuple $(v, \lambda, \alpha, \delta, \beta, \zeta, \xi, \bar{t})$ satisfying a (SSNE) condition and it is analogous to Property (P3) for a tuple $(v, \lambda, \alpha, \delta, \beta, \bar{t})$ satisfying a (SNE) condition.

Property (SP3). Given $\mu, \nu > 0$, it is possible to re-design a tuple $(v, \lambda, \alpha, \delta, \beta, \zeta, \xi, \bar{t})$ satisfying a (SSNE) condition into a new tuple $(\tilde{v}, \lambda, \tilde{\alpha}, \tilde{\delta}, \tilde{\beta}, \tilde{\zeta}, \tilde{\xi}, \bar{t})$, with $\tilde{\beta}(s) := \nu s^\mu$, still satisfying a (SSNE) condition (the proof is found in section B of the appendix).

VI. DESIGN OF \mathcal{H} -FILTERS

In view of the introductory explanation of section IV, our second task is to identify, under general assumptions on A, C, \mathcal{F} and \mathcal{H} , a canonical class of filters (which, in what follows, we refer to as \mathcal{H} -filters) for systems of the form (6) which throughout this section we address as:

$$\begin{aligned} \dot{\mathbf{e}}_t &= A\mathbf{e}_t + \hat{\Psi}_t + \mathcal{F}(\mathbf{e}_t + \hat{\mathbf{X}}_t, \mathbf{d}_t) \\ \mathbf{w}_t &= C\mathbf{e}_t + \mathcal{H}(\mathbf{e}_t + \hat{\mathbf{X}}_t, \mathbf{d}_t), \end{aligned} \quad (34)$$

where $\hat{\Psi}_t$ and $\hat{\mathbf{X}}_t$ are exogenous known inputs. These \mathcal{H} -filters are adaptively tuned by a \mathcal{N} -filter, which will be assumed at hand together with its associated tuple $(v, \lambda, \alpha, \delta, \beta, \zeta, \xi, \bar{t})$ satisfying a (SSNE) condition. An elegant mathematical tool for designing the tuning action of a \mathcal{N} -filter on a \mathcal{H} -filter is the notion of family of dilations.

¹We introduce this extra condition for reducing the design complexity and we reassure the reader that considering a linear $\zeta \in \mathcal{K}$ is not a technical issue (this will be addressed in a forthcoming paper).

A. Family of dilations

Let Λ be a family of solutions of the following matrix differential equation

$$\frac{d\Lambda}{ds}(s) = \frac{W\Lambda(s)}{s}, \quad s > 0, \quad (35)$$

where $W \in \mathbb{R}^{n \times n}$ is any symmetric² matrix with all eigenvalues in \mathbb{C}^- . As well-known, $\Lambda(s) := e^{\ln(s)W}\lambda$, for $s > 0$ and $\lambda > 0$, is such parametrized family of solutions. Notice also that, given $\delta \in \mathcal{K}_\infty^1$ and $\mu > 0$, it is always possible to rescale Λ as

$$\Lambda[\delta, \mu](s) := \Lambda(\delta(s))\mu, \quad s > 0, \quad (36)$$

which is a parametrized family of solutions of

$$\frac{d\Lambda[\delta, \mu]}{ds}(s) = \frac{1}{\delta(s)} \frac{d\delta}{ds}(s) W \Lambda[\delta, \mu](s), \quad s > 0. \quad (37)$$

We call $\Lambda^{-1}[\delta, \mu]$ (the inverse of $\Lambda[\delta, \mu]$) a *family of dilations* if $\Lambda[\delta, \mu]$ is a parametrized family of (37) and we denote by \mathcal{G} the set of families of dilations (in more general frameworks families of dilations are introduced in [18] and [13]). We can define an equivalence relation in \mathcal{G} as follows: for $\Lambda^{-1}[\delta_1, \mu_1], \Lambda^{-1}[\delta_2, \mu_2] \in \mathcal{G}$ we say that $\Lambda^{-1}[\delta_1, \mu_1] \sim \Lambda^{-1}[\delta_2, \mu_2]$ if there exist $\mu > 0$ and $\delta \in \mathcal{K}_\infty^1$ such that $\Lambda[\delta_2, \mu_2](s) \equiv \mu \Lambda[\delta_1, \mu_1](\delta(s))$. In this case we say that $\Lambda^{-1}[\delta_2, \mu_2] \in \mathcal{G}$ *rescales* $\Lambda^{-1}[\delta_1, \mu_1] \in \mathcal{G}$ by $(\mu, \delta) \in \mathbb{R}_> \times \mathcal{K}_\infty^1$. In what follows we omit, whenever no ambiguity shows up, the term between square brackets in $\Lambda[\delta, \mu]$ (resp. $\Lambda^{-1}[\delta, \mu]$).

Remark 6.1: Family of dilations $\Lambda^{-1} \in \mathcal{G}$ of the form $\Lambda^{-1}(s) = e^{-\ln(s)(\alpha W)}\lambda$, with $\alpha > 0$, are all equivalent. Therefore, we can rescale a family of dilations simply by rescaling the eigenvalues of W by $\alpha > 0$. \square

B. Discussion of the main assumptions on A, C, \mathcal{F} and \mathcal{H}

The notion of family of dilations is naturally associated with the notion of (generalized) homogeneity, which we assume here on the vector field $f(e) := Ae$ and the map $h(e) := Ce$. **(HL)** (*Generalized homogeneous linearization*). There exist $\Lambda^{-1} \in \mathcal{G}$, $\gamma_A \in \mathcal{K}_0$, $\gamma_C \in \mathcal{K}_\infty$ such that for all $s > 0$

$$\frac{\Lambda(s)A\Lambda^{-1}(s)}{\gamma_A(s)} = A, \quad \frac{C\Lambda^{-1}(s)}{\gamma_C(s)} = C. \quad (38)$$

Remark 6.2: While $\gamma_C \in \mathcal{K}_\infty$ is a reasonable assumption (on account of the second condition in (38) and the fact that Λ^{-1} contains increasing modes: see example 6.1), a more natural condition for γ_A would be³ $\gamma_A \in \mathcal{K}_0 \cup \mathcal{L}_\infty$ (according to how the decreasing modes of Λ interact, through A , with the increasing modes of Λ^{-1} in the first condition in (38)). Under this regard, we will say only that the case $\gamma_A \in \mathcal{L}_\infty$ requires a more sophisticated design technique for the \mathcal{N} -filter based

²This assumption is made only for simplifying our analysis and can be directly omitted.

³With $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $C = (1 \ 0)$, it is easy to see that $\gamma_A(s) = s \in \mathcal{K}_0$, $\gamma_C(s) = s \in \mathcal{K}_\infty$ with $\Lambda(s) := \text{diag}\{s^{-1}, s^{-2}\}$ but $\gamma_A(s) = s^{-1} \in \mathcal{L}_\infty$, $\gamma_C(s) = s^2 \in \mathcal{K}_\infty$ with $\Lambda(s) := \text{diag}\{s^{-2}, s^{-1}\}$. On the other hand, $\gamma_A(s) = 1 \in \mathcal{K}_0$, $\gamma_C(s) = s \in \mathcal{K}_\infty$ with $\Lambda(s) := \text{diag}\{s^{-1}, s^{-1}\}$.

on a suitable time rescaling, which goes beyond the scope of this paper.

Assumption (HL) can be relaxed by requiring homogeneity only in the ∞ -limit as follows:

$$(\mathbf{HL}_\infty) : \lim_{s \rightarrow +\infty} \frac{\Lambda(s)A\Lambda^{-1}(s)}{\gamma_A(s)} = A_0, \quad \lim_{s \rightarrow +\infty} \frac{C\Lambda^{-1}(s)}{\gamma_C(s)} = C_0,$$

for some $A_0 \in \mathbb{R}^{n \times n}$ and $C_0 \in \mathbb{R}^{p \times n}$, at the price of a sufficiently large choice of $\hat{\gamma} > 0$ in the state norm estimate (3). However, with (\mathbf{HL}_∞) we can capture many pairs (C, A) which are left out by⁴ (HL). \square

Next, we introduce similar (generalized) homogeneity assumptions for dominating the nonlinearities of (34) via the homogeneous linearization. The idea of homogeneously dominating the nonlinearities using the linearization comes from previous works as for example, to cite a few, [17] and, more recently, [1]. Here, we introduce conditions not dependent on specific system's structures and less restrictive, also taking into account the presence of noise. Recall from the notation section that for given $F \in \mathcal{P}_>^m$, $G \in \mathcal{P}_\geq^n$ and $S \in \mathbb{R}^{m \times n}$ we write $\|S\|_{(F,G)} \leq \gamma$ by meaning that $\|Sw\|_F \leq \gamma\|w\|_G$ for all $w \in \mathbb{R}^n$ or equivalently $S^\top FS \leq \gamma^2 G$.

(HD $_\infty$) (Generalized homogeneous domination at ∞). For all $(x, d, s) \in \mathbb{R}^n \times D \times [1, +\infty)$

$$\left\| \Lambda(s) \frac{\partial \mathcal{F}(w, 0)}{\partial w} \Big|_{w=\Lambda^{-1}(s)x} \Lambda^{-1}(s) \right\|_{(F_A(x), \Sigma_A(x))} \leq \gamma_A(s) \quad (39)$$

$$\left\| \frac{\partial \mathcal{H}(w, 0)}{\partial w} \Big|_{w=\Lambda^{-1}(s)x} \Lambda^{-1}(s) \right\|_{(H_C(x), \Xi_C(x))} \leq \gamma_C(s) \quad (40)$$

$$\left\| \Lambda(s) \frac{\partial \mathcal{F}(w, d)}{\partial d} \Big|_{w=\Lambda^{-1}(s)x} \right\|_{(F_d(x,d), \Sigma_d(x,d))} \leq \gamma_A(s) \gamma_d(s) \quad (41)$$

$$\left\| \frac{\partial \mathcal{H}(w, d)}{\partial d} \Big|_{w=\Lambda^{-1}(s)x} \right\|_{(H_d(x,d), \Xi_d(x,d))} \leq \gamma_C(s) \gamma_d(s) \quad (42)$$

with $\gamma_d \in \mathcal{L}_\infty^1$ and continuous functions $F_A : \mathbb{R}^n \rightarrow \mathcal{P}_>^n$, $\Sigma_A : \mathbb{R}^n \rightarrow \mathcal{P}_\geq^n$, $F_d : \mathbb{R}^n \times D \rightarrow \mathcal{P}_>^n$, $\Sigma_d : \mathbb{R}^n \times D \rightarrow \mathcal{P}_\geq^n$, $H_C : \mathbb{R}^n \rightarrow \mathcal{P}_\geq^p$, $\Xi_C : \mathbb{R}^n \rightarrow \mathcal{P}_\geq^p$, $H_d : \mathbb{R}^n \times D \rightarrow \mathcal{P}_\geq^p$ and $\Xi_d : \mathbb{R}^n \times D \rightarrow \mathcal{P}_\geq^p$ with $\Sigma_A(0) = 0$ and $\Xi_C(0) = 0$.

By setting

$$\begin{aligned} \gamma(s) &:= \{\gamma_A(s), \gamma_C(s), \gamma_d(s)\}, \\ V(x) &:= \{F_A(x), \Sigma_A(x), \Xi_C(x), H_C(x)\}, \\ Z(x, d) &:= \{F_d(x, d), \Sigma_d(x, d), H_d(x, d), \Xi_d(x, d)\}, \end{aligned}$$

we also say that $(\Lambda^{-1}, \gamma, V, Z)$ satisfy a (HL)+(HD $_\infty$) condition. The matrices $F_A, F_d, H_C, H_d, \Sigma_A, \Sigma_d, \Xi_C$ and Ξ_d are weight matrices which can be properly selected (see example 6.1 below).

Remark 6.3: (Globally Lipschitz \mathcal{F} and \mathcal{H}). If \mathcal{F} and \mathcal{H} are globally Lipschitz, (HL), (\mathbf{HL}_∞) , (HD $_\infty$) are all satisfied with the simple choice

$$\Lambda(s) = (1/s)I, \quad \gamma(s) = \{1, s, 1/s\}$$

and any choice of the weights consistent with (39)-(42). If $\frac{\partial \mathcal{F}(x, 0)}{\partial x} \equiv 0$ (resp. $\frac{\partial \mathcal{H}(x, 0)}{\partial x} \equiv 0$) we can take $\Sigma_A \equiv 0$ (resp.

⁴For instance, $C = C_0 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$, $A = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}$, $A_0 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ with $\Lambda(s) = \begin{pmatrix} s^{-1} & 0 \\ 0 & s^{-2} \end{pmatrix}$ and $\gamma_A(s) = \gamma_C(s) = s$.

$\Xi_C \equiv 0$) with arbitrary F_A (resp. H_C) and if $\frac{\partial \mathcal{F}(x, d)}{\partial d} \equiv 0$ (resp. $\frac{\partial \mathcal{H}(x, d)}{\partial d} \equiv 0$) we can take $\Sigma_d \equiv 0$ (resp. $\Xi_d \equiv 0$) with arbitrary F_d (resp. H_d). \square

Remark 6.4: (Invariance of (HL)+(HD $_\infty$) condition under rescaling of Λ^{-1}). An important property of the (HL)+(HD $_\infty$) condition, extensively used in designing the chain of \mathcal{K} -filters, is that this condition is invariant under rescaling (or re-parametrization) of Λ^{-1} . In other words, by re-parametrizing Λ^{-1} as $\Lambda^{-1} \circ \delta$, $\delta \in \mathcal{K}_\infty^1$, condition (HL)+(HD $_\infty$) remains valid by re-parametrizing also γ as $\gamma \circ \delta$. To be more precise, let \mathcal{E} be the set of tuples $(\Lambda^{-1}, \gamma, V, Z)$ satisfying a (HL)+(HD $_\infty$) condition. The equivalence relation $\Lambda^{-1}[\delta_1, \mu_1] \sim \Lambda^{-1}[\delta_2, \mu_2]$ in \mathcal{G} induces, via (HL) and (HD $_\infty$), an equivalence relation $(\Lambda^{-1}, \gamma, V, Z)[\delta_1, \mu_1] \sim (\Lambda^{-1}, \gamma, V, Z)[\delta_2, \mu_2]$ in \mathcal{E} as follows: there exist $\delta \in \mathcal{K}_\infty^1$ and $\mu > 0$ such that

$$\begin{aligned} \gamma[\delta_2, \mu_2](s) &\equiv \left\{ \gamma_A[\delta_1, \mu_1](\delta(s)), \frac{1}{\mu} \gamma_C[\delta_1, \mu_1](\delta(s)), \right. \\ &\quad \left. \mu \gamma_d[\delta_1, \mu_1](\delta(s)) \right\}, \\ V[\delta_2, \mu_2](x) &\equiv V[\delta_1, \mu_1] \left(\frac{x}{\mu} \right), \\ Z[\delta_2, \mu_2](x, d) &\equiv Z[\delta_1, \mu_1] \left(\frac{x}{\mu}, d \right). \end{aligned} \quad (43)$$

Therefore, (HL)+(HD $_\infty$) are invariant under rescalings in \mathcal{G} in the following sense: if $\Lambda^{-1}[\delta_1, \mu_1] \sim \Lambda^{-1}[\delta_2, \mu_2]$ in \mathcal{G} then $(\Lambda^{-1}, \gamma, V, Z)[\delta_1, \mu_1] \sim (\Lambda^{-1}, \gamma, V, Z)[\delta_2, \mu_2]$ in \mathcal{E} . \square

Remark 6.5: After a normalization of (γ, V, Z) we can assume

$$\gamma_A(s) \geq 1, \quad \forall s \geq 1, \quad (44)$$

and for some $\gamma_{d,0} < 0$

$$\gamma_d(s) = s^{-|\gamma_{d,0}|}, \quad \forall s \geq 1 \quad (45)$$

(the proof is found in section C of the appendix). \square

Example 6.1: Consider

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \end{pmatrix},$$

$$\mathcal{F}(x, d) = (0, -x_2 x_1^2 + x_1 d)^\top, \quad \mathcal{H}(x, d) = 0. \quad (46)$$

It is easy to see that $(\Lambda^{-1}, \gamma, V, Z)$ satisfy a (HL)+(HD $_\infty$) condition, with $\Lambda^{-1}(s) := e^{-\ln(s)W}$, $W := \text{diag}\{-\tau, -3\tau\}$ and $\gamma(s) = \{\gamma_A(s), \gamma_C(s), \gamma_d(s)\} = \{s^{2\tau}, s^\tau, s^{-3\tau}\}$ for any $\tau > 0$. The weights (V, Z) are chosen as follows:

$$\begin{aligned} F_A(x) &= (1 + \varphi_A(x))^{-1} \text{diag}\{\varepsilon, 1\}, \\ \Sigma_A(x) &= \varphi_A(x)(1 + \varphi_A(x))^{-1} I, \\ F_d(x) &= (1 + |x_1|)^{-1} \text{diag}\{\varepsilon, 1\}, \quad \Sigma_d(x) = |x_1|(1 + |x_1|)^{-1}, \end{aligned} \quad (47)$$

with any $\varepsilon > 0$ and $\varphi_A(x) := 5(x_1^2 x_2^2 + x_1^4)$. The remaining weights H_C, H_d, Ξ_C and Ξ_d are completely arbitrary since $\mathcal{H} \equiv 0$ (remark 6.3): for instance, $H_C = H_d := \mu > 0$ with any $\mu > 0$, $\Xi_C = 0$ and $\Xi_d = 0$. \triangleleft

We follow up with some assumptions on \mathbf{e}_t and the exogenous input $\hat{\mathbf{X}}_t$. Let $\hat{\mathbf{v}}_t$ be the output of the \mathcal{N} -filter associated with the tuple $(v, \lambda, \alpha, \delta, \beta, \zeta, \xi, \bar{t})$ satisfying a (SSNE) condition.

(B_∞) (Asymptotic bounds). There exist $e_{d_\infty}, \widehat{X}_{d_\infty} > 0$ (depending only on d_∞) such that

$$\limsup_{t \rightarrow +\infty} \|\Lambda(\widehat{\mathbf{v}}_t + \widehat{\gamma})\mathbf{e}_t\| < e_{d_\infty}, \quad (48)$$

$$\limsup_{t \rightarrow +\infty} \|\Lambda(\widehat{\mathbf{v}}_t + \widehat{\gamma})\widehat{\mathbf{X}}_t\| < \widehat{X}_{d_\infty}. \quad (49)$$

We also say that $(e_{d_\infty}, \widehat{X}_{d_\infty}, \Lambda^{-1})$ satisfy a **(B_∞)** condition on (34).

Remark 6.6: The asymptotic bounds (48) and (49) are in general guaranteed, when designing the chain of \mathcal{K} -filters (see recursive algorithm in section VII-C), by the choice of the parameters of the previous \mathcal{K} -filters in the chain. We state (48) and (49) as an assumption so that to relate directly the numbers $e_{d_\infty}, \widehat{X}_{d_\infty}$ with the design parameters of the \mathcal{K} -filter (see the filter equations (54), (55) below). Moreover, when (34) coincides with (2) (i.e. $\widehat{\Psi}_t \equiv 0$ and $\widehat{\mathbf{X}}_t \equiv 0$), which corresponds to the situation in which we are designing the first \mathcal{K} -filter in the chain, **(B_∞)** is not assumed *a priori* but guaranteed by design of the \mathcal{N} -filter (see Theorem 6.2 below). \square

Finally, we introduce a stability margin condition for guaranteeing estimation error convergence.

(SM) (Stability margin). There exist

$$\lambda_\pi > \bar{\lambda}_\pi := 4(|\gamma_{d,0}| + |\lambda_{min}^W|)(\zeta + \xi(d_\infty) + \delta(d_\infty)) \quad (50)$$

and $\Pi \in \mathcal{P}_>^n$ such that:

$$\text{Lyap}(\Pi, W) := \Pi W + W \Pi < 0, \quad (51)$$

$$\begin{aligned} \text{Ric}(\Pi, \lambda_\pi, c) &:= \Pi(A + \lambda_\pi I) + (A + \lambda_\pi I)^\top \Pi \\ &+ \Pi B(c)\Pi - C^\top R^{-1}(c)C \leq -M(c) \end{aligned} \quad (52)$$

with $c := e_{d_\infty} + \widehat{X}_{d_\infty}$ and

$$\begin{aligned} M(s) &:= \sup_{\|x_1\| \leq 2ns} (\Sigma_A(x_1) + \Xi_C(x_1)) \\ B(s) &:= \sup_{\substack{\|x_1\| \leq 2ns \\ \|x_2\| \leq ns \\ \|d\| \leq d_\infty}} (F_A^{-1}(x_1) + F_d^{-1}(x_2, d)) \\ R(s) &:= \sup_{\substack{\|x_1\| \leq 2ns \\ \|x_2\| \leq ns \\ \|d\| \leq d_\infty}} (H_C^{-1}(x_1) + H_d^{-1}(x_2, d)). \end{aligned} \quad (53)$$

We also say that the tuple $(e_{d_\infty}, \widehat{X}_{d_\infty}, \Lambda^{-1}, \gamma, V, Z)$ satisfy a **(SM)** condition with some (Π, λ_π) . In (53) as $\sup_{(x,d) \in \mathcal{S}} T(x,d)$ we mean and use any matrix \bar{T} such that $T(x,d) \leq \bar{T}$ for all $(x,d) \in \mathcal{S}$.

Remark 6.7: (Design of Π). Notice that since $\sigma(W) \subset \mathbb{C}^-$, there always exists $\Pi \in \mathcal{P}_>^n$ for which $\text{Lyap}(\Pi, W) < 0$. We require in assumption **(SM)** that $\Pi \in \mathcal{P}_>^n$ at the same time satisfies $\text{Ric}(\Pi, \lambda_\pi, c) \leq -M(c)$. The matrix $M(c)$ represents a guaranteed *margin* for the Riccati equation (52) and depends on the weights Σ_A and Ξ_C .

Few more comments on the solution of the Lyapunov-Riccati inequalities (51) and (52). For each $\lambda_\pi > 0$ and $c > 0$ for which $(C, A + \lambda_\pi I)$ is detectable and $(A + \lambda_\pi I, B_0(c))$ is controllable with $B_0(c)B_0^\top(c) = B(c)$, it is well-known that $\text{Ric}(\Pi, \lambda_\pi, c) \leq 0$ has always a solution $\Pi \in \mathcal{P}_>^n$. Moreover, if (C, A) is in observability canonical form, $(A, B_0(c))$ in controllability canonical form and W is diagonal, we can prove the existence of $\Pi \in \mathcal{P}_>^n$ satisfying at the same time $\text{Lyap}(\Pi, W) < 0$ and $\text{Ric}(\Pi, \lambda_\pi, c) \leq 0$. In addition, the

margin $M(c)$ is even 0 when $\frac{\partial \mathcal{F}(x,0)}{\partial x} \equiv 0$ and $\frac{\partial \mathcal{H}(x,0)}{\partial x} \equiv 0$ (i.e. multiplicative noise).

In more canonical EKF-based approaches as in [7] a differential state-dependent Riccati equation is assumed, by requiring uniform lower and upper bounds for its time-varying solution. This is too demanding in a nonlinear context. In our framework, homogeneity domination assumptions **(HD_∞)** with asymptotic bounds **(B_∞)** allow us to ask only for the common solution $\Pi \in \mathcal{P}^n$ of a pair of algebraic Lyapunov-Riccati inequalities. \square

Example 6.1 (cont'd). With the choice of the weights (V, Z) , parametrized by $\varepsilon, \mu > 0$, in example 6.1, for any given $e_{d_\infty}, \widehat{X}_{d_\infty} > 0$, we can always find $\varepsilon, \mu > 0$ for which $(e_{d_\infty}, \widehat{X}_{d_\infty}, \Lambda^{-1}, \gamma, V, Z)$ satisfy a **(SM)** condition for some (Π, λ_π) . In practice, in this example we can guarantee any margin M and handle arbitrary asymptotic values $e_{d_\infty}, \widehat{X}_{d_\infty} > 0$ in **(B_∞)**. \triangleleft

C. Canonical \mathcal{K} -filters and main results

A \mathcal{K} -filter for (34) has the following parametric form:

$$\begin{aligned} \dot{\widehat{\mathbf{e}}}_t &= A\widehat{\mathbf{e}}_t + \widehat{\Psi}_t + \widehat{\mathcal{F}}(\widehat{\mathbf{e}}_t, \widehat{\mathbf{X}}_t, \Lambda(\widehat{\mathbf{v}}_t + \widehat{\gamma})) + \mathcal{K}(\widehat{\mathbf{v}}_t + \widehat{\gamma})(\mathbf{w}_t - \widehat{\mathbf{w}}_t), \\ \widehat{\mathbf{w}}_t &= C\widehat{\mathbf{e}}_t + \widehat{\mathcal{H}}(\widehat{\mathbf{e}}_t, \widehat{\mathbf{X}}_t, \Lambda(\widehat{\mathbf{v}}_t + \widehat{\gamma})), \end{aligned} \quad (54)$$

$$\begin{aligned} \mathcal{K}(z) &:= \frac{\gamma A(z)}{(\gamma_C(z))^2} \mathcal{P}^{-1}(z) R^{-1}(e_{d_\infty} + \widehat{X}_{d_\infty}) C^\top, \\ \mathcal{P}(z) &:= \Lambda^\top(z) \Pi \Lambda(z), \end{aligned} \quad (55)$$

$$\begin{aligned} \widehat{\mathcal{F}}(\widehat{\mathbf{e}}, \widehat{X}, \Lambda) &:= \mathcal{F}\left(\Lambda^{-1} \text{sat}_{e_{d_\infty}}(\Lambda \widehat{\mathbf{e}}) + \Lambda^{-1} \text{sat}_{\widehat{X}_{d_\infty}}(\Lambda \widehat{X}), 0\right), \\ \widehat{\mathcal{H}}(\widehat{\mathbf{e}}, \widehat{X}, \Lambda) &:= \mathcal{H}\left(\Lambda^{-1} \text{sat}_{e_{d_\infty}}(\Lambda \widehat{\mathbf{e}}) + \Lambda^{-1} \text{sat}_{\widehat{X}_{d_\infty}}(\Lambda \widehat{X}), 0\right). \end{aligned}$$

All the parameters of (54)-(55) are directly obtained from the given assumptions **(HL)**, **(HD_∞)**, **(B_∞)** and **(SM)**.

Remark 6.8: The reason for which a \mathcal{K} -filter has a Kalman-type feature lies in the definition of its gain matrix \mathcal{K} in (55). For linear systems (34), $\frac{\partial \mathcal{F}}{\partial x} \equiv 0$, $\frac{\partial \mathcal{H}}{\partial x} \equiv 0$ and $\frac{\partial \mathcal{F}}{\partial d}$ and $\frac{\partial \mathcal{H}}{\partial d}$ are constant. Moreover, as mentioned in remark 6.3, we can take $\Lambda(s) = (1/s)I$, $\gamma(s) = \{1, s, 1/s\}$ and, assuming invertible $B := \frac{\partial F}{\partial d} \left(\frac{\partial F}{\partial d} \right)^\top$ and $R := \frac{\partial H}{\partial d} \left(\frac{\partial H}{\partial d} \right)^\top$, the weights (V, Z) can be selected so as $\Sigma_A + \Xi_C = 0$, $F_A^{-1} + F_d^{-1} = B$ and $H_d^{-1} + H_C^{-1} = R$. From definition (55) we have

$$\mathcal{K} = \Pi^{-1} R^{-1} C^\top \quad (56)$$

where, from (52), $\Pi \in \mathcal{P}_>^n$ is a solution of the Riccati equation

$$\Pi(A + \lambda_\pi I) + (A + \lambda_\pi I)^\top \Pi + \Pi B \Pi - C^\top R^{-1} C = 0. \quad (57)$$

Notice that since $\Lambda(s) = (1/s)I$ then $W = -I$ and (51) trivially holds with the same $\Pi \in \mathcal{P}_>^n$ solution of (57). The matrices B and R are representative of the state and, respectively, measurement noise covariances in a stochastic setting and the Riccati equation (57) (neglecting the term $\lambda_\pi I$ which guarantees a certain non-zero margin) is the one associated to a steady state Kalman filter with steady state error covariance Π^{-1} and Kalman gain (56). It is interesting from (52) and (53) to see how the nonlinearities enter the picture through the weights F_A^{-1} and H_C^{-1} by affecting the margin M . \square

We are ready to state the first result of this section (the proof has been postponed to the appendix), which points out conditions for the estimation error convergence of a \mathcal{K} -filter (54)-(55) for (34).

Theorem 6.1: *Assume (SSNE), (HL), (HD_∞), (B_∞) and (SM). Along the solutions of (34), (54), (17) we have*

$$\limsup_{t \rightarrow +\infty} \frac{\|\Lambda(\hat{\mathbf{v}}_t + \hat{\gamma})(\mathbf{e}_t - \hat{\mathbf{e}}_t)\|}{\gamma_d(\hat{\mathbf{v}}_t + \hat{\gamma})} < c_{d_\infty} := \frac{2\rho_\infty(e_{d_\infty} + \hat{X}_{d_\infty}, d_\infty)}{\sqrt{\lambda_\pi \lambda_{\min}^\Pi}}, \quad (58)$$

where $\rho_\infty(r, s) := \sup_{\substack{\|d\| \leq s, \\ \|x\| \leq nr}} \|d\|_{\Sigma_d(x, d) + \Xi_d(x, d)}$. *If in addition*

$$|\gamma_{d,0}| \geq |\lambda_{\min}^W|, \quad (59)$$

then

$$\limsup_{t \rightarrow +\infty} \|\mathbf{e}_t - \hat{\mathbf{e}}_t\| < \gamma_\infty c_{d_\infty} \quad (60)$$

where $\gamma_\infty := \sqrt{\sup_{s \geq 1} \gamma_d^2(s) s^{2|\lambda_{\min}^W|}}$.

Remark 6.9: Theorem 6.1 presents two asymptotic bounds for the estimation error. The tighter one is presented in (60) under the additional condition (59), i.e. the decreasing rate of γ_d is at least $|\lambda_{\min}^W|$. As it results from (58), the \mathcal{K} -filter (54) asymptotically steers the (weighted) estimation error norm $\frac{\|\Lambda(\hat{\mathbf{v}}_t + \hat{\gamma})(\mathbf{e}_t - \hat{\mathbf{e}}_t)\|}{\gamma_d(\hat{\mathbf{v}}_t + \hat{\gamma})}$ inside the interval $[0, c_{d_\infty}]$. However, we cannot say the same for the estimation error norm $\|\mathbf{e}_t - \hat{\mathbf{e}}_t\|$ unless (59) holds true. Without this condition, at least one eigenvalue of the matrix $W + |\gamma_{d,0}|I$ is negative and let denote by $\bar{\lambda}$ its minimum eigenvalue: $\bar{\lambda} := \min_j \{\lambda_j \in \sigma(W + |\gamma_{d,0}|I)\} < 0$. On account of (45) it follows from (58) that

$$\|\mathbf{e}_t - \hat{\mathbf{e}}_t\| \leq (\hat{\mathbf{v}}_t + \hat{\gamma})^{|\bar{\lambda}|} c_{d_\infty} \quad (61)$$

for $t \geq t_0$ (depending on the initial error). The bound (61) on $\|\mathbf{e}_t - \hat{\mathbf{e}}_t\|$ becomes large with $\hat{\mathbf{v}}_t^{|\bar{\lambda}|}$ and can be directly compared with the tighter bound (60). The smaller the size of $|\bar{\lambda}|$ (or, which is the same, the excess $\Delta := |\lambda_{\min}^W| - |\gamma_{d,0}|$), the smaller the bound on $\|\mathbf{e}_t - \hat{\mathbf{e}}_t\|$ in (61). Notice also that, in our context, $\hat{\mathbf{v}}_t$ grows unbounded in time if \mathbf{x}_t does (since $\hat{\mathbf{v}}_t$ is a state norm estimate) and the error $\mathbf{e}_t - \hat{\mathbf{e}}_t$ will grow unbounded according to (61). On the other hand, if \mathbf{x}_t is bounded in time, $\hat{\mathbf{v}}_t$ as well as $\mathbf{e}_t - \hat{\mathbf{e}}_t$ will be bounded in time according to (61), which can be recast in the form (60) with a γ_∞ depending on the upper bound of $\hat{\mathbf{v}}_t$. For these reasons, the upper bound (60), which depends only on d_∞ , is much more satisfactory than the upper bound (61), which depends on d_∞ but also on the large variations of $\hat{\mathbf{v}}_t$.

The rate condition (59) guarantees the tighter asymptotic bound (60) on the estimation error norm. In example 6.1, since $\Delta = 0$ and the rate condition (59) is met, the estimation error norm satisfies an asymptotic bound (60) and the \mathcal{K} -filter (54) has good error performances. However, an extra additive measurement disturbance in example 6.1 (i.e. $\mathcal{H}(x, d) = d \neq 0$) has a deteriorating effect on the error performances: indeed, with $\mathcal{H}(x, d) = d$ and A, C, \mathcal{F} as in example 6.1, (HL)+(HD_∞) is satisfied with the same Λ^{-1} , $\gamma_A(s)$ and $\gamma_C(s)$ but $\gamma_d(s) = s^{-\tau}$ (i.e. slower decrease rate). Since now $\Delta = 2\tau > 0$ we can only obtain a weaker

asymptotic bound (61) and worse error performances. A simple explanation of this performance deterioration lies in the *high-gain* feature of the \mathcal{K} -filter (54) in example 6.1: indeed, in this case $\mathcal{H}(z) = \text{diag}(z^{2\tau}, z^{4\tau})\Pi^{-1}(1, 0)^T$ (definition (55)) and an extra additive measurement disturbance has the effect of deteriorating the error performances of the high-gain \mathcal{K} -filter (54) (compare with semiglobal high-gain observers for feedback linearizable systems: [12], [2], [19], [11], [8]). The asymptotic bound (58) corresponds in our framework to deteriorated error performances and the excess Δ can be taken as a quantitative margin for error performances evaluation. This motivates in the next section the design of chains of \mathcal{K} -filters in which the excess Δ is progressively reduced and the error performances of the \mathcal{K} -filters tend to an ideal situation characterized by the tighter asymptotic bound (60) (in which the excess is 0).

The asymptotic bound (60) corresponds to good error performances. If in addition $\mathcal{F}(0, d) \equiv 0$ and $\mathcal{H}(0, d) \equiv 0$ (multiplicative state and measurement disturbances) the noise sensitivity can be reduced to the same extent as $c := e_{d_\infty} + \hat{X}_{d_\infty}$: indeed, if $\mathcal{F}(0, d) \equiv 0$ and $\mathcal{H}(0, d) \equiv 0$ then $\Sigma_d(0, d) \equiv 0$ and $\Xi_d(0, d) \equiv 0$ and $\rho_\infty(0, s) = 0$ for all $s \geq 0$ so that c_∞ in (60) can be made small if $e_{d_\infty}, \hat{X}_{d_\infty}$ are small. Notice how the asymptotic bound (60) depends only on d_∞ (since $e_{d_\infty} + \hat{X}_{d_\infty}$ and, therefore, c_{d_∞} does by assumption (B_∞)): in general, estimation error bounds depend explicitly on the initial state in global frameworks ([4], [21]) or implicitly through the invariant compact set containing the state trajectories in semi-global frameworks ([12], [19], [11]). Therefore, the error performances (in term of robustness and noise sensitivity) of the \mathcal{K} -filter (54) in case of excess $\Delta = 0$ are good even for, and actually independent of large state initial conditions. \square A parallel result to Theorem 6.1 is relative to the system (34) when it coincides with (2) (i.e. $\hat{\Psi}_t \equiv 0$ and $\hat{\mathbf{X}}_t \equiv 0$). As mentioned in the introductory section IV, this corresponds to the case in which we design the first \mathcal{K} -filter of the chain. The canonical form of the \mathcal{K} -filter is simpler in this case:

$$\begin{aligned} \dot{\hat{\mathbf{x}}}_t &= A\hat{\mathbf{x}}_t + \widehat{\mathcal{F}}(\hat{\mathbf{x}}_t, \Lambda(\hat{\mathbf{v}}_t + \hat{\gamma})) + \mathcal{K}(\hat{\mathbf{v}}_t + \hat{\gamma})(\mathbf{y}_t - \hat{\mathbf{y}}_t), \\ \hat{\mathbf{y}}_t &= C\hat{\mathbf{x}}_t + \widehat{\mathcal{H}}(\hat{\mathbf{x}}_t, \Lambda(\hat{\mathbf{v}}_t + \hat{\gamma})), \end{aligned} \quad (62)$$

$$\begin{aligned} \mathcal{K}(z) &:= \frac{\gamma_A(z)}{(\gamma_C(z))^2} \mathcal{P}^{-1}(z) R^{-1}(c) C^T, \quad \mathcal{P}(z) := \Lambda^T(z) \Pi \Lambda(z), \\ \widehat{\mathcal{F}}(\hat{x}, \Lambda) &:= \mathcal{F}(\Lambda^{-1} \text{sat}_c(\Lambda \hat{x}), 0), \\ \widehat{\mathcal{H}}(\hat{x}, \Lambda) &:= \mathcal{H}(\Lambda^{-1} \text{sat}_c(\Lambda \hat{x}), 0) \end{aligned} \quad (63)$$

for $c > 0$. As we will see, the asymptotic bounds (B_∞) are not *a priori* assumed here (but guaranteed by the design of the \mathcal{N} -filter) and only an arbitrarily small margin M must be guaranteed in (SM). The proof of the next result has been moved to section D of the appendix.

Theorem 6.2: *Assume (SSNE), (HL), (HD_∞) and the existence of $c > 0$ such that $(c, 0, \Lambda^{-1}, \gamma, V, Z)$ satisfy a (SM) condition with some (Π, λ_π) . There exists a tuple $(\tilde{v}, \tilde{\lambda}, \tilde{\alpha}, \tilde{\delta}, \tilde{\beta}, \tilde{\zeta}, \tilde{\xi}, \tilde{t})$, satisfying a (SSNE) condition with associated \mathcal{N} -filter*

$$\dot{\hat{\mathbf{v}}}_t = -\tilde{\lambda} \hat{\mathbf{v}}_t + \tilde{\alpha}(\|\mathbf{y}_t\|) + \tilde{\delta}(d_\infty) \quad (64)$$

such that along the solutions of (2), (62) and (64) we have

$$\limsup_{t \rightarrow +\infty} \frac{\|\Lambda(\hat{\mathbf{v}}_t + \hat{\gamma})(\mathbf{x}_t - \hat{\mathbf{x}}_t)\|}{\gamma_d(\hat{\mathbf{v}}_t + \hat{\gamma})} < c_{d_\infty} := \frac{2\rho_\infty(c, d_\infty)}{\sqrt{\lambda_\pi \lambda_{min}^\Pi}}, \quad (65)$$

with ρ_∞ as in Theorem 6.1. If, in addition, (59) holds true then

$$\limsup_{t \rightarrow +\infty} \|\mathbf{x}_t - \hat{\mathbf{x}}_t\| < \gamma_\infty c_{d_\infty} \quad (66)$$

with γ_∞ as in Theorem 6.1.

Remark 6.10: (Design of the \mathcal{N} -filter). From the proof of Theorem 6.2 (in section E of the appendix) it turns out that, for any given $\Delta_0 \in (0, 1)$, the tuple $(v, \lambda, \alpha, \delta, \beta, \zeta, \xi, \bar{t})$ satisfying a (SSNE) condition can be always transformed into a new tuple $(\tilde{v}, \tilde{\lambda}, \tilde{\alpha}, \tilde{\delta}, \tilde{\beta}, \tilde{\zeta}, \tilde{\xi}, \tilde{t})$, with

$$\tilde{\beta}(s) := \left(\frac{s}{c}\right)^{\frac{1}{\Delta_0 |\lambda_{max}^W|}}, \quad (67)$$

still satisfying a (SSNE) condition with associated \mathcal{N} -filter (64) such that for all $\hat{\gamma} > 1$ and $\bar{\Delta} \in (\Delta_0, 1]$

$$\limsup_{t \rightarrow +\infty} \|\Lambda((\hat{\mathbf{v}}_t + \hat{\gamma})^{\bar{\Delta}})\mathbf{x}_t\| < c. \quad (68)$$

In other words, the \mathcal{N} -filter (64) is designed in such a way to dominate \mathbf{x}_t under the action of Λ .

Remark 6.11: (Design of the parameter c). In practice, the parameter $c > 0$ such that $(c, 0, \Lambda^{-1}, \gamma, V, Z)$ satisfy a (SM) condition with some (Π, λ_π) can be obtained in two steps as follows. First, determine $(\bar{\Pi}, \bar{\lambda}_\pi)$ (if any) such that $\text{Ric}(\bar{\Pi}, W, \bar{\lambda}_\pi, 0) < 0$ and $\bar{\Pi}W + W\bar{\Pi} < 0$. The first inequality can be solved under simple detectability and controllability conditions on A, C , and $B(0)$ (see for instance remark 6.7). Secondly, using the fact that $M(0) = 0$ (since $F_A(0) = 0$ and $\Sigma_C(0) = 0$ by assumption (HD_∞)) and by continuity, it is easy to find $c > 0$ such that $(c, 0, \Lambda^{-1}, \gamma, V, Z)$ satisfy a (SM) condition with $(\Pi, \lambda_\pi) = (\bar{\Pi}, \bar{\lambda}_\pi)$. Notice how (SM) is satisfied with an arbitrarily small margin $M(c)$, corresponding to arbitrarily small values of $c > 0$.

It is worth noticing that the above design procedure for finding (Π, λ_π) and c such that $(c, 0, \Lambda^{-1}, \gamma, V, Z)$ satisfy a (SM) condition cannot be used to find (Π, λ_π) such that $(e_{d_\infty}, \hat{X}_{d_\infty}, \Lambda^{-1}, \gamma, V, Z)$ satisfying a (SM) condition (as in theorem 6.1): indeed, in the first case c is obtained *a posteriori* from Π using continuity (and can be chosen arbitrarily small) while in the second case Π is computed *a posteriori* from $c := e_{d_\infty} + \hat{X}_{d_\infty}$ which is not arbitrarily small. However, the magnitude of $c := e_{d_\infty} + \hat{X}_{d_\infty}$ can be kept small by increasing $\hat{\gamma}$ or choosing smaller $e_{d_\infty}, \hat{X}_{d_\infty} > 0$ (see remarks after (72)). \square

VII. CHAINS OF \mathcal{K} -FILTERS FOR SEQUENTIAL ERROR PERFORMANCE OPTIMIZATION

The aim of the next sections is to show that by sequentially reducing the excess Δ for each \mathcal{K} -filter, we obtain tighter bounds (58) while tending to the ideal situation which corresponds to the tightest bound (60). The reduction of the excess Δ can be obtained by rescaling the family of dilations Λ^{-1} and, on the other hand, the rescaling of Λ^{-1} corresponds to rescale the eigenvalues of W in such a way to reduce their reciprocal distance: as a matter of fact, the ideal situation is

the one for which the eigenvalues of W are all equal, a typical situation when \mathcal{F} and \mathcal{H} are globally Lipschitz (see remark 6.3). Before dipping into the details of sequential processing, we introduce a comparison criterion for error performances evaluation.

A. A comparison criterion for state estimates

The next definition gives a simple criterion to compare two different estimates $\hat{\mathbf{X}}_t^{(k)}$, $k = 1, 2$, of \mathbf{x}_t .

Definition 7.1: Let $\hat{\mathbf{v}}_t$ be the output of a \mathcal{N} -filter, with $\hat{\mathbf{X}}_t^{(1)}$ and $\hat{\mathbf{X}}_t^{(2)}$ estimates of \mathbf{x}_t such that $\limsup_{t \rightarrow +\infty} \frac{\|\mathbf{x}_t - \hat{\mathbf{X}}_t^{(k)}\|}{\omega^{(k)}(\hat{\mathbf{v}}_t)} < c^{(k)}(d_\infty)$, $k = 1, 2$, for some $\omega^{(k)} \in \mathcal{K}_0$ and $c^{(k)} \in \mathcal{K}_\infty$. If $\limsup_{t \rightarrow +\infty} \frac{\omega^{(2)}(\hat{\mathbf{v}}_t)c^{(2)}(d_\infty)}{\omega^{(1)}(\hat{\mathbf{v}}_t)c^{(1)}(d_\infty)} := r < 1$, we say that $\hat{\mathbf{X}}_t^{(2)}$ outperforms $\hat{\mathbf{X}}_t^{(1)}$ by $(1 - r) (\times 100) \%$.

In other words, $\hat{\mathbf{X}}_t^{(2)}$ has better error performances than $\hat{\mathbf{X}}_t^{(1)}$ if it corresponds to a tighter (in percentage) asymptotic bound on the estimation error norm. Obviously, the fact that $\hat{\mathbf{X}}_t^{(2)}$ outperforms $\hat{\mathbf{X}}_t^{(1)}$ by $(1 - r) (\times 100) \%$ does not guarantee that the ratio between the actual values of $\|\mathbf{x}_t - \hat{\mathbf{X}}_t^{(2)}\|$ and $\|\mathbf{x}_t - \hat{\mathbf{X}}_t^{(1)}\|$ be (even asymptotically) less than r . In our framework the comparison is made on some conservative asymptotic bounds on $\|\mathbf{x}_t - \hat{\mathbf{X}}_t^{(2)}\|$ and, respectively, $\|\mathbf{x}_t - \hat{\mathbf{X}}_t^{(1)}\|$, but by reducing progressively the ratio between these conservative bounds it is likely that also the ratio between the actual values of $\|\mathbf{x}_t - \hat{\mathbf{X}}_t^{(2)}\|$ and $\|\mathbf{x}_t - \hat{\mathbf{X}}_t^{(1)}\|$ will ultimately decrease. The comparison criterion will be used to compare the estimates associated to two consecutive \mathcal{K} -filters in the chain.

B. Motivations and outline of sequential processing

For explaining what we mean by sequential processing and how it is implemented in practice, we assume to have at hand a \mathcal{N} -filter together with a \mathcal{K} -filter for (2). For this reason, we assume:

- (A1) a tuple $(v, \lambda, \alpha, \delta, \beta, \zeta, \xi, \bar{t})$ satisfying a (SSNE) condition,
- (A2) a tuple $(\Lambda^{-1}, \gamma, V, Z)$ satisfying a (HL)+(HD $_\infty$) condition,
- (A3) the existence of $c > 0$ such that $(c, 0, \Lambda^{-1}, \gamma, V, Z)$ satisfy a (SM) condition with some (Π, λ_π) .

As it follows from (SP3), the tuple $(v, \lambda, \alpha, \delta, \beta, \zeta, \xi, \bar{t})$ and the associated \mathcal{N} -filter can be re-designed so that (68) holds for any given $\Delta_0 \in (0, \frac{\Delta}{|\lambda_{min}^W|}]$ and for all $\hat{\gamma} > 1$ and $\bar{\Delta} \in (\Delta_0, 1]$. On the other hand, under the given assumptions (A2)-(A3), the \mathcal{K} -filter is designed as pointed out in Theorem 6.2, with $\hat{\mathbf{x}}_t$ being the estimate of \mathbf{x}_t and $\mathbf{e}_t := \mathbf{x}_t - \hat{\mathbf{x}}_t$ being the corresponding estimation error.

Assume the excess $\Delta := |\lambda_{min}^W| - |\gamma_{d,0}| > 0$ (otherwise, as explained in remark 6.9, the error performances of the \mathcal{K} -filter can be considered satisfactory). From Theorem 6.2 we get the asymptotic bound (65) on \mathbf{e}_t :

$$\limsup_{t \rightarrow +\infty} \frac{\|\Lambda(\hat{\mathbf{v}}_t + \hat{\gamma})\mathbf{e}_t\|}{\gamma_d(\hat{\mathbf{v}}_t + \hat{\gamma})} < c_{d_\infty} := \frac{2\rho_\infty(c, d_\infty)}{\sqrt{\lambda_\pi \lambda_{min}^\Pi}}. \quad (69)$$

On account of (69) with (45), it follows that ⁵

$$\limsup_{t \rightarrow +\infty} \|\bar{\Lambda}(\hat{\mathbf{v}}_t + \hat{\gamma})\mathbf{e}_t\| < e_{d_\infty} := c_{d_\infty} \quad (70)$$

where $\bar{\Lambda}^{-1}$ is the rescaling of Λ^{-1} by $(1, \bar{\delta})$ with $\bar{\delta}(s) := s\bar{\Delta}$ and $\bar{\Delta} \in [\frac{\Delta}{|\lambda_{min}^W|}, 1)$. As observed in remark 6.1 this corresponds, equivalently, to rescale the eigenvalues of W by $\bar{\Delta}$, i.e. $\bar{\Lambda}(s) = e^{\ln(s)(\bar{W})}\lambda$ where $\bar{W} := \bar{\Delta}W$. With (70) at hand and using Theorem 6.1, the idea is to design a second \mathcal{K} -filter, sequentially connected to the first \mathcal{K} -filter, for estimating the error \mathbf{e}_t , described by the equations:

$$\begin{aligned} \dot{\mathbf{e}}_t &= \mathbf{A}\mathbf{e}_t + \Psi_t + \mathcal{F}(\mathbf{e}_t + \hat{\mathbf{x}}_t, \mathbf{d}_t), \\ \mathbf{w}_t &= \mathbf{C}\mathbf{e}_t + \mathcal{H}(\mathbf{e}_t + \hat{\mathbf{x}}_t, \mathbf{d}_t), \end{aligned} \quad (71)$$

with $\mathbf{w}_t := \mathbf{y}_t - \mathbf{C}\hat{\mathbf{x}}_t$, $\Psi_t := \mathbf{A}\hat{\mathbf{x}}_t - \dot{\hat{\mathbf{x}}}_t$. For applying Theorem 6.1 to (71), which has the general form (34), we need to prove that all the assumptions of Theorem 6.1 are met. First of all, by assumption (A2) and invariance of (HL)+(HD_∞) under rescalings of Λ^{-1} (remark 6.3), $(\bar{\Lambda}^{-1}, \bar{\gamma}, V, Z)$ satisfy a (HL)+(HD_∞) condition with $\bar{\gamma}(s) := \gamma(\bar{\delta}(s))$.

Since $\bar{\Lambda}(s) = \Lambda(s\bar{\Delta})$ and $\hat{\mathbf{x}}_t = \mathbf{x}_t - \hat{\mathbf{e}}_t$, from (68) and (70)

$$\limsup_{t \rightarrow +\infty} \|\bar{\Lambda}(\hat{\mathbf{v}}_t + \hat{\gamma})\hat{\mathbf{x}}_t\| < \hat{X}_{d_\infty} := c + e_{d_\infty}. \quad (72)$$

With (70) and (72) at hand, it follows straightforwardly that $(e_{d_\infty}, \hat{X}_{d_\infty}, \bar{\Lambda}^{-1})$ satisfy a (B_∞) condition on (71).

Finally, we need to prove that $(e_{d_\infty}, \hat{X}_{d_\infty}, \bar{\Lambda}^{-1}, \bar{\gamma}, V, Z)$ satisfy a (SM) condition with some $(\bar{\Pi}, \bar{\lambda}_\pi)$. In a first scenario, no matter the values of $e_{d_\infty}, \hat{X}_{d_\infty}$ are, $(e_{d_\infty}, \hat{X}_{d_\infty}, \bar{\Lambda}^{-1}, \bar{\gamma}, V, Z)$ indeed satisfy a (SM) condition (this is the case, for instance, of example 6.1, in which any guaranteed margin M is achievable). In a second scenario, $(e_{d_\infty}, \hat{X}_{d_\infty}, \bar{\Lambda}^{-1}, \bar{\gamma}, V, Z)$ cannot satisfy a (SM) for any $(\bar{\Pi}, \bar{\lambda}_\pi)$. In this case, we may resort to the following heuristics. Since also $(c, 0, \bar{\Lambda}^{-1}, \bar{\gamma}, V, Z)$ satisfy a (SM) condition with (Π, λ_π) , we can try smaller values of $e_{d_\infty} < c_{d_\infty}$ and $\hat{X}_{d_\infty} < c + c_{d_\infty}$ as close as possible to 0 and, respectively, c so that $(e_{d_\infty}, \hat{X}_{d_\infty}, \bar{\Lambda}^{-1}, \bar{\gamma}, V, Z)$ satisfy a (SM) condition with some $(\bar{\Pi}, \bar{\lambda}_\pi)$ close to (Π, λ_π) . The possibility that these smaller values of $e_{d_\infty} < c_{d_\infty}$ and $\hat{X}_{d_\infty} < c + c_{d_\infty}$ still satisfy a (B_∞) condition is accounted for by the fact that in practice the estimation error norm is asymptotically smaller than the state or its estimate norm and since $\liminf_{t \rightarrow +\infty} \hat{\mathbf{v}}_t \gg 1$. Moreover, we can also reduce the magnitude of $e_{d_\infty} + \hat{X}_{d_\infty}$ by increasing $\hat{\gamma}$ and, therefore, decreasing the magnitude of $\bar{\Lambda}(\hat{\mathbf{v}}_t + \hat{\gamma})\mathbf{e}_t$ and $\bar{\Lambda}(\hat{\mathbf{v}}_t + \hat{\gamma})\hat{\mathbf{x}}_t$ in (70) and, respectively, (72).

With the second \mathcal{K} -filter designed on (71) according to Theorem 6.1 and with the estimate $\hat{\mathbf{e}}_t$ of \mathbf{e}_t at hand, we obtain

⁵We notice that since W is symmetric there exists orthonormal T such that $W = T^T W_D T$ with $W_D = \text{diag}\{\lambda_i \in \sigma(W), i = 1, \dots, n\}$. Using standard properties of matrix exponentials and $\|T\| = \|T^T\| = 1$ and recalling that $\Delta = |\lambda_{min}^W| - |\gamma_{d,0}|$, since $\bar{\Delta} \in [\frac{\Delta}{|\lambda_{min}^W|}, 1)$ and by (45), for all $z \geq 1$ and $e \in \mathbb{R}^n$ we get $\|\bar{\Lambda}(z)e\| \leq \frac{1}{\gamma_d(z)} \|e^{\ln(z)((\bar{\Delta}-1)W - |\gamma_{d,0}|I)}\| \|\Lambda(z)e\| \leq \frac{z^{\Delta - \bar{\Delta}|\lambda_{min}^W|}}{\gamma_d(z)} \|\Lambda(z)e\| \leq \frac{\|\Lambda(z)e\|}{\gamma_d(z)}$. By virtue of this, from (69) we get (70).

from Theorem 6.1 the following asymptotic bound on the error $\bar{\mathbf{e}}_t := \mathbf{e}_t - \hat{\mathbf{e}}_t$, analogue to the bound (69) on \mathbf{e}_t :

$$\limsup_{t \rightarrow +\infty} \frac{\|\bar{\Lambda}(\hat{\mathbf{v}}_t + \hat{\gamma})\bar{\mathbf{e}}_t\|}{\bar{\gamma}_d(\hat{\mathbf{v}}_t + \hat{\gamma})} < \bar{c}_{d_\infty} := \frac{\rho_\infty(e_{d_\infty} + \hat{X}_{d_\infty}, d_\infty)}{\sqrt{\lambda_\pi \lambda_{min}^{\bar{\Pi}}}}. \quad (73)$$

It is important to stress the fact that \bar{c}_{d_∞} depends only on d_∞ (and this will remain so in all subsequent iterations), since e_{d_∞} and \hat{X}_{d_∞} depend by construction only on d_∞ .

The current state estimate $\hat{\mathbf{X}}_t^{(1)} := \hat{\mathbf{x}}_t$ is updated by the second \mathcal{K} -filter as $\hat{\mathbf{X}}_t^{(2)} := \hat{\mathbf{X}}_t^{(1)} + \hat{\mathbf{e}}_t$. By comparison of the asymptotic bounds on $\mathbf{e}_t = \mathbf{x}_t - \hat{\mathbf{X}}_t^{(1)}$ in (69) and, respectively, $\bar{\mathbf{e}}_t = \mathbf{x}_t - \hat{\mathbf{X}}_t^{(2)}$ in (73) we establish by which amount in percentage $\hat{\mathbf{X}}_t^{(2)}$ outperforms $\hat{\mathbf{X}}_t^{(1)}$ (according to the criterion 7.1). The other \mathcal{K} -filters in the chain are likewise designed.

C. A recursive algorithm for designing the chain of \mathcal{K} -filters

Along the general lines highlighted in the previous section, we give here the details of a recursive algorithm for the design of a chain of \mathcal{K} -filters. As mentioned at the beginning of section VII-B, we assume (A1)-(A3) with excess $\Delta > 0$.

(I) (Initialization: Design of the \mathcal{N} -filter and the first \mathcal{K} -filter in the chain). Let $\Delta_0 \in (0, (\frac{\Delta}{|\lambda_{min}^W|})^N]$, where $N \in \mathbb{N}$ is an estimate of the maximum number of iterations to be performed by the algorithm. Let $\hat{\mathbf{v}}_t$ be the output of a \mathcal{N} -filter, designed so that (68) holds for all $\bar{\Delta} \in (\Delta_0, 1]$ and $\hat{\gamma} > 1$. Let

$$\begin{aligned} (\Lambda^{(1)-1}, \gamma^{(1)}, V^{(1)}, Z^{(1)}) &:= (\Lambda^{-1}, \gamma, V, Z), \\ c_{d_\infty}^{(2)} &:= \frac{2\rho_\infty(c, d_\infty)}{\sqrt{\lambda_\pi \lambda_{min}^{\bar{\Pi}}}}, \Pi^{(1)} := \Pi, W^{(1)} := W, \end{aligned}$$

Moreover, $\gamma_d(s) = s^{-|\gamma_{d,0}|}$ and $\gamma_A(s) \geq 1$ for all $s \geq 1$, for some $\gamma_{d,0} < 0$ (remark 6.5). Set $\gamma_{d,0}^{(1)} := \gamma_{d,0}$ and $\Delta^{(1)} := \Delta$.

Let $\hat{\mathbf{x}}_t$ be the output of the \mathcal{K} -filter designed as pointed out in theorem 6.2 and $\mathbf{e}_t := \mathbf{x}_t - \hat{\mathbf{x}}_t$ be the corresponding state estimation error. Set $\mathbf{e}_t^{(2)} := \mathbf{e}_t$ and $\hat{\mathbf{X}}_t^{(1)} := \hat{\mathbf{x}}_t$, the initial state estimate of \mathbf{x}_t . From Theorem 6.2 we have

$$\limsup_{t \rightarrow +\infty} \frac{\|\Lambda^{(1)}(\hat{\mathbf{v}}_t + \hat{\gamma})\mathbf{e}_t^{(2)}\|}{\gamma_d^{(1)}(\hat{\mathbf{v}}_t + \hat{\gamma})} < c_{d_\infty}^{(2)}. \quad (74)$$

Set $k = 2$.

(II) (Design of the k -th \mathcal{K} -filter in the chain). Consider the rescaled family of dilations

$$\Lambda^{(k)-1}(s) := \Lambda^{(k-1)-1}(\delta^{(k)}(s)) \quad (75)$$

where

$$\delta^{(k)}(s) := s^{\bar{\Delta}^{(k)}}, \bar{\Delta}^{(k)} \in \left[\frac{\Delta^{(k-1)}}{|\lambda_{min}^{W^{(k-1)}}|}, 1 \right). \quad (76)$$

Also, let $e_{d_\infty}^{(k)}, \hat{X}_{d_\infty}^{(k)} > 0$ be such that

$$\limsup_{t \rightarrow +\infty} \|\Lambda^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})\mathbf{e}_t^{(k)}\| < e_{d_\infty}^{(k)}, \quad (77)$$

$$\limsup_{t \rightarrow +\infty} \|\Lambda^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})\hat{\mathbf{X}}_t^{(k-1)}\| < \hat{X}_{d_\infty}^{(k)}. \quad (78)$$

Define the new tuple

$$(\gamma^{(k)}, V^{(k)}, Z^{(k)}) := (\gamma^{(k-1)}(\delta^{(k)}), V^{(k-1)}, Z^{(k-1)}) \quad (79)$$

and $\Pi^{(k)} \in \mathcal{P}_>^n$ be such that $(e_{d_\infty}^{(k)}, \hat{X}_{d_\infty}^{(k)}, \gamma^{(k)}, V^{(k)}, Z^{(k)})$ satisfy a (SM) condition with $(\Pi^{(k)}, \lambda_\pi)$. From Theorem 6.1 we get the $\mathcal{K}^{(k)}$ -filter and the asymptotic bound

$$\limsup_{t \rightarrow +\infty} \frac{\|\Lambda^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})\mathbf{e}_t^{(k+1)}\|}{\gamma_d^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})} < c_\infty^{(k+1)}(d_\infty) \quad (80)$$

where $\mathbf{e}_t^{(k+1)} := \mathbf{e}_t^{(k)} - \hat{\mathbf{e}}_t^{(k)}$ and $\hat{\mathbf{e}}_t^{(k)}$ is the output of the $\mathcal{K}^{(k)}$ -filter.

(III) (State estimate update). Set $\gamma_{d,0}^{(k)} := \gamma_{d,0}^{(k-1)}\bar{\Delta}^{(k)}$, $W^{(k)} := W^{(k-1)}\bar{\Delta}^{(k)}$, $\Delta^{(k)} := |\lambda_{min}^{W^{(k)}}| - |\gamma_{d,0}^{(k)}| = \Delta^{(k-1)}\bar{\Delta}^{(k)}$ and

$$c_\infty^{(k+1)}(d_\infty) := \frac{2\rho_\infty(e_{d_\infty}^{(k)} + \hat{X}_{d_\infty}^{(k)}, d_\infty)}{\sqrt{\lambda_\pi \lambda_{min}^{\Pi^{(k)}}}}. \quad (81)$$

Update the estimate $\hat{\mathbf{X}}_t^{(k-1)}$ of \mathbf{x}_t as: $\hat{\mathbf{X}}_t^{(k)} := \hat{\mathbf{X}}_t^{(k-1)} + \hat{\mathbf{e}}_t^{(k)}$.

(IV) (Error performances evaluation). If

$$\frac{c_{d_\infty}^{(k+1)}/c_{d_\infty}^{(k)}}{(\liminf_{t \rightarrow +\infty} \hat{\mathbf{v}}_t + \hat{\gamma})^{\Delta^{(k-1)}(1 - \bar{\Delta}^{(k)})}} := r^{(k)} < 1 \quad (82)$$

then $\hat{\mathbf{X}}_t^{(k)}$ outperforms $\hat{\mathbf{X}}_t^{(k-1)}$ by $(1 - r^{(k)}) \times 100\%$.

(V) $k \rightarrow k + 1$ and goto (II).

Remark 7.1: Some explanatory remarks are in order for the design parameters $(e_{d_\infty}^{(k)}, \hat{X}_{d_\infty}^{(k)})$ in (77), (78) and for the outperformance condition (82). To clear up any confusion, we stress the fact that the existence of $(e_{d_\infty}^{(k)}, \hat{X}_{d_\infty}^{(k)})$ satisfying (77), (78) is always guaranteed by the $(k - 1)$ -th iteration. First of all, since from the $(k - 1)$ -th iteration we have

$$\limsup_{t \rightarrow +\infty} \frac{\|\Lambda^{(k-1)}(\hat{\mathbf{v}}_t + \hat{\gamma})\mathbf{e}_t^{(k)}\|}{\gamma_d^{(k-1)}(\hat{\mathbf{v}}_t + \hat{\gamma})} < c_{d_\infty}^{(k)}, \quad (83)$$

by definition of $\Lambda^{(k-1)}$ in (75) it is guaranteed at the k -th iteration that

$$\limsup_{t \rightarrow +\infty} \|\Lambda^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})\mathbf{e}_t^{(k)}\| < c_{d_\infty}^{(k)} \quad (84)$$

(in the same way we concluded (70) from (69) in section VII-B). If, for any given values of $e_{d_\infty}^{(k)}$ and $\hat{X}_{d_\infty}^{(k)}$, $(e_{d_\infty}^{(k)}, \hat{X}_{d_\infty}^{(k)}, \gamma^{(k)}, V^{(k)}, Z^{(k)})$ satisfy a (SM) condition (as in example 6.1), in view of (84) we pick any $e_{d_\infty}^{(k)} \geq c_{d_\infty}^{(k)}$ in (77). If not, we should notice from the $(k - 1)$ -th iteration and the definition of $\Lambda^{(k)}$ and $\gamma^{(k)}$ that $(e_{d_\infty}^{(k-1)}, \hat{X}_{d_\infty}^{(k-1)}, \gamma^{(k)}, V^{(k)}, Z^{(k)})$ satisfy a (SM) condition with $(\Pi^{(k-1)}, \lambda_\pi)$. In order to have $(e_{d_\infty}^{(k)}, \hat{X}_{d_\infty}^{(k)})$ as close as possible to $(e_{d_\infty}^{(k-1)}, \hat{X}_{d_\infty}^{(k-1)})$ so that $(e_{d_\infty}^{(k)}, \hat{X}_{d_\infty}^{(k)}, \gamma^{(k)}, V^{(k)}, Z^{(k)})$ is likely to satisfy a (SM) condition for some $\Pi^{(k)}$ close to $\Pi^{(k-1)}$, we rather pick values of $e_{d_\infty}^{(k)} \ll c_{d_\infty}^{(k)}$. As already accounted for in section VII-B, this heuristic approach, although not *a priori* guaranteeing the validity of (77) with $e_{d_\infty}^{(k)} \ll c_{d_\infty}^{(k)}$, anyway has reasonable chances of success in view of the fact that in practice $c_{d_\infty}^{(k)} \gg 1$ and $\limsup_{t \rightarrow +\infty} \|\Lambda^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})\mathbf{e}_t^{(k)}\| \ll 1$.

Moreover, as already noticed, increasing $\hat{\gamma}$ may be beneficial in decreasing the magnitude of $e_{d_\infty}^{(k)}$ by the corresponding decrease in magnitude of $\Lambda^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})\mathbf{e}_t^{(k)}$ in (77).

Once (77) is satisfied with a suitable choice of $e_{d_\infty}^{(k)} > 0$, directly from $\hat{\mathbf{X}}_t^{(k-1)} = -\mathbf{e}_t^{(k)} + \mathbf{x}_t$ and (68) we get

$$\limsup_{t \rightarrow +\infty} \|\Lambda^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})\hat{\mathbf{X}}_t^{(k-1)}\| < e_{d_\infty}^{(k)} + c. \quad (85)$$

If, for any given values of $e_{d_\infty}^{(k)}$ and $\hat{X}_{d_\infty}^{(k)}$, $(e_{d_\infty}^{(k)}, \hat{X}_{d_\infty}^{(k)}, \gamma^{(k)}, V^{(k)}, Z^{(k)})$ satisfy a (SM) condition (as in example 6.1), in view of (85) we pick any $\hat{X}_{d_\infty}^{(k)} \geq e_{d_\infty}^{(k)} + c$ in (78). If not, in order to have $e_{d_\infty}^{(k)} + \hat{X}_{d_\infty}^{(k)}$ as close as possible to $e_{d_\infty}^{(k-1)} + \hat{X}_{d_\infty}^{(k-1)}$ so that $(e_{d_\infty}^{(k)}, \hat{X}_{d_\infty}^{(k)}, \gamma^{(k)}, V^{(k)}, Z^{(k)})$ is likely to satisfy a (SM) condition for some $\Pi^{(k)}$ close to $\Pi^{(k-1)}$, we rather pick values of $\hat{X}_{d_\infty}^{(k)} < e_{d_\infty}^{(k)} + c$ as close as possible to c but still with no *a priori* guarantee on the validity of (78). Moreover, as already noticed, increasing $\hat{\gamma}$ may be beneficial in decreasing the magnitude of $\hat{X}_{d_\infty}^{(k)}$ by the corresponding decrease in magnitude of $\Lambda^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})\hat{\mathbf{X}}_t^{(k-1)}$ in (78).

Finally, few comments on (82). This condition is the result of applying the comparison criterion 7.1 to the estimates $\hat{\mathbf{X}}_t^{(k)}$ and, respectively, $\hat{\mathbf{X}}_t^{(k-1)}$ with

$$\omega^{(k)}(\hat{\mathbf{v}}_t) := \frac{\gamma_d^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})c_{d_\infty}^{(k+1)}}{\sqrt{\lambda_{min}^{\Lambda^{(k)\top}(\hat{\mathbf{v}}_t + \hat{\gamma})\Lambda^{(k)}(\hat{\mathbf{v}}_t + \hat{\gamma})}}} = (\hat{\mathbf{v}}_t + \hat{\gamma})^{\Delta^{(k)}} c_{d_\infty}^{(k+1)}$$

and using $\limsup_{t \rightarrow +\infty} (1/\hat{\mathbf{v}}_t) \leq 1/(\liminf_{t \rightarrow +\infty} \hat{\mathbf{v}}_t)$. From the $(k - 1)$ -th iteration and the choice of $(e_{d_\infty}^{(k)}, \hat{X}_{d_\infty}^{(k)})$ and $\Pi^{(k)}$ we have no *a priori* guarantee that $r^{(k)} < 1$. However, since in practice $\liminf_{t \rightarrow +\infty} \hat{\mathbf{v}}_t \gg 1$ it is reasonable to expect $r^{(k)} < 1$ for at least a number of iterations after which either $c_{d_\infty}^{(k+1)}/c_{d_\infty}^{(k)} \gg 1$ or $\Delta^{(k-1)}(1 - \bar{\Delta}^{(k)}) \ll 1$. This is even more evident in unstable systems (2) for which $\liminf_{t \rightarrow +\infty} \|\mathbf{x}_t\| = +\infty$ and, consistently with the properties of the state norm estimator, $\liminf_{t \rightarrow +\infty} \hat{\mathbf{v}}_t = +\infty$ so that we always have $r^{(k)} = 0$. On the other hand, if $\frac{\partial \mathcal{F}(x,d)}{\partial d}$ and $\frac{\partial \mathcal{H}(x,d)}{\partial d}$ are uniformly bounded (for instance, when the disturbances are additive) then $\rho_\infty(s_1, r) = \rho_\infty(s_2, r)$ for all $s_1, s_2, r \geq 0$ and the numerator term in (82) is equal to $\sqrt{\frac{\lambda_{min}^{\Pi^{(k-1)}}}{\lambda_{min}^{\Pi^{(k)}}}}$. If, in addition, $e_{d_\infty}^{(k)} + \hat{X}_{d_\infty}^{(k)} \approx e_{d_\infty}^{(k-1)} + \hat{X}_{d_\infty}^{(k-1)}$ then $\Pi^{(k)} \approx \Pi^{(k-1)}$ and $\sqrt{\frac{\lambda_{min}^{\Pi^{(k-1)}}}{\lambda_{min}^{\Pi^{(k)}}}} \approx 1$: in this case condition (82) is satisfied with $r^{(k)} \ll 1$ since in practice $\liminf_{t \rightarrow +\infty} \hat{\mathbf{v}}_t \gg 1$. \square

D. Simulations

Consider the noisy Duffing oscillator

$$\dot{\mathbf{x}}_{1,t} = \mathbf{x}_{2,t}, \quad \dot{\mathbf{x}}_{2,t} = \mathbf{x}_{1,t} - \mathbf{x}_{1,t}^3 + \mathbf{d}_{1,t}, \quad \mathbf{y}_t = \mathbf{x}_{1,t} + \mathbf{d}_{2,t} \quad (86)$$

with initial state $\mathbf{x}_0 = (5, -8)^\top$, $\mathbf{d}_{2,t}$ a sinusoidal disturbance with frequency 20 and amplitude 2 while $\mathbf{d}_{1,t}$ a sinusoidal disturbance with frequency 1 and amplitude 1.

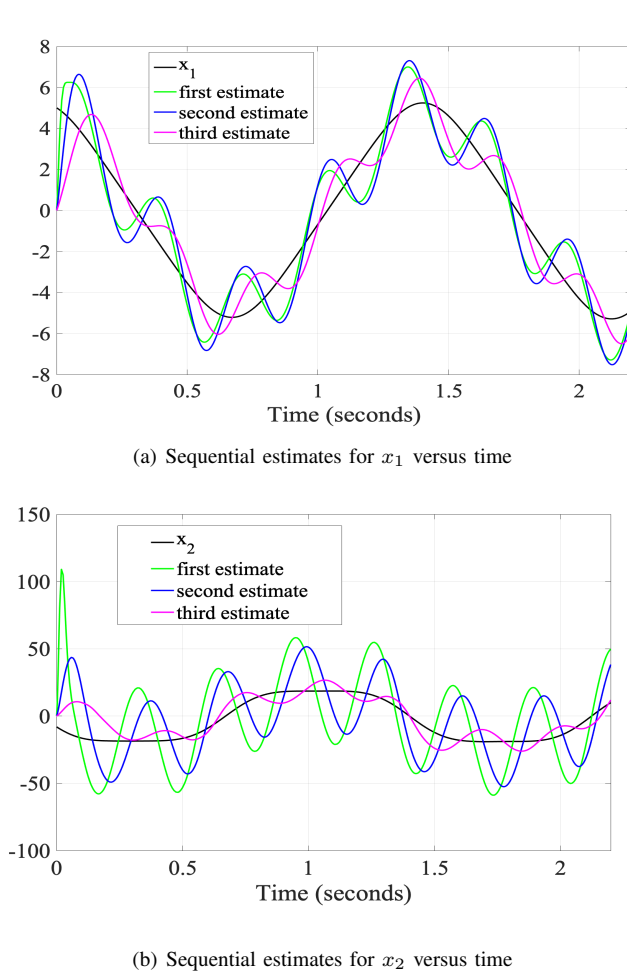


Figure 1. State \mathbf{x}_t (black line), first estimate $\hat{\mathbf{X}}_t^{(1)}$ (light green line), second sequential estimate $\hat{\mathbf{X}}_t^{(2)}$ (light blue line), third sequential estimate $\hat{\mathbf{X}}_t^{(3)}$ (magenta line).

Assumption (A1) is satisfied for (86) and a \mathcal{N} -filter has been designed according to the lines of section VII-C: $\hat{\mathbf{v}}_t = -\hat{\mathbf{v}}_t + 8[5 + \mathbf{y}_t^2 + 16(\mathbf{y}_t^2 + 4)(\mathbf{y}_t^4 + 16)]$. Assumptions (A2) and (A3) have been satisfied with $W = \text{diag}\{-0.4, -0.8\}$ and $\gamma_A(s) = s^{0.4}$, $\gamma_C(s) = s^{0.4}$, $\gamma_d(s) = s^{-0.4}$. Notice that the excess $\Delta := |\lambda_{\min}^W| - |\gamma_d, 0| = 0.4 > 0$ (mainly, for the presence of the additive measurement disturbance \mathbf{d}_2). Moreover,

$$\Pi = \begin{pmatrix} 2.5 & -1 \\ -1 & 0.5 \end{pmatrix}, \lambda_\pi = 0.1, c = 0.07, \varrho_\infty = 15, c_{d_\infty}^{(2)} = 44.$$

The state \mathbf{x}_t and the estimate $\hat{\mathbf{X}}_t^{(1)} := \hat{\mathbf{x}}_t$, computed by the first \mathcal{K} -filter designed as in step (I) of the algorithm VII-C, are shown versus time in Figs. 1a) and 1b) (black and light green lines).

A second \mathcal{K} -filter has been designed as in step (II) of the algorithm VII-C for estimating $\mathbf{e}_t^{(2)} := \mathbf{x}_t - \hat{\mathbf{X}}_t^{(1)}$. We used $\bar{\Delta}^{(2)} = 0.5$ (therefore, the excess is reduced by a factor 0.5), $e_{d_\infty}^{(2)} = 0.04 \in (0, c_{d_\infty}^{(2)}) = (0, 44]$ and $\hat{X}_{d_\infty}^{(2)} = 0.7 \geq e_{d_\infty}^{(2)} + c = 0.11$. Moreover, since $e_{d_\infty}^{(2)} + \hat{X}_{d_\infty}^{(2)} = 0.15 \approx c$ we used $\Pi^{(2)} \approx \Pi^{(1)} = \Pi$ so that $c_{d_\infty}^{(3)} \approx c_{d_\infty}^{(2)}$. Using step (IV) of

the algorithm VII-C, we figure out that $\hat{\mathbf{X}}_t^{(2)}$ outperforms $\hat{\mathbf{X}}_t^{(1)}$ by $\approx 75\%$ (since $\liminf_{t \rightarrow +\infty} \hat{\mathbf{v}}_t > 10^3$). With $\hat{\mathbf{e}}_t^{(2)}$ being the estimate of $\mathbf{e}_t^{(2)}$, the state \mathbf{x}_t and its second sequential estimate $\hat{\mathbf{X}}_t^{(2)} := \hat{\mathbf{X}}_t^{(1)} + \hat{\mathbf{e}}_t^{(2)}$ (updated as in step (III) of the algorithm VII-C) are shown versus time in Figs. 1a) and 1b) (black and blue lines).

A third \mathcal{K} -filter has been designed as in step (II) of the algorithm VII-C $\mathbf{e}_t^{(3)} := \mathbf{x}_t - \hat{\mathbf{X}}_t^{(2)}$. We used $\bar{\Delta}^{(3)} = 0.5$ (once again the excess is reduced by a factor 0.5), $e_{d_\infty}^{(3)} = 0.4 \in (0, c_{d_\infty}^{(3)}) = (0, 44)$ and $\hat{X}_{d_\infty}^{(3)} = 2.3 > e_{d_\infty}^{(3)} + c \approx 0.47$. Moreover, since $e_{d_\infty}^{(3)} + \hat{X}_{d_\infty}^{(3)} = 2.7 > e_{d_\infty}^{(2)} + \hat{X}_{d_\infty}^{(2)}$ we used some $\Pi^{(3)} > \Pi^{(2)}$ in (SM). Also in this case, using step (IV) of the algorithm VII-C, we figure out that $\hat{\mathbf{X}}_t^{(3)}$ outperforms $\hat{\mathbf{X}}_t^{(2)}$ by $\approx 75\%$. With $\hat{\mathbf{e}}_t^{(3)}$ being the estimate of $\mathbf{e}_t^{(3)}$, the state \mathbf{x}_t and its third sequential estimate $\hat{\mathbf{X}}_t^{(3)} := \hat{\mathbf{X}}_t^{(2)} + \hat{\mathbf{e}}_t^{(3)}$ (updated as in step (III) of the algorithm VII-C) are shown versus time in Figs. 1a) and 1b) (black and magenta lines).

VIII. CONCLUSIONS

We have proposed a framework for designing chains of \mathcal{K} -filters with global convergence properties, robustness and performing noise sensitivity, adaptively driven by a \mathcal{N} -filter. The state estimate of each \mathcal{K} -filter is sequentially processed by the next \mathcal{K} -filter in the chain optimizing the noise sensitivity by reducing the asymptotic estimation error bound. A recursive algorithm has been given for designing the chain of filters and future work will be devoted to improving few critical issues discussed in section 7.1.

APPENDIX

A. Proof of (P3)

Let $(v, \lambda, \alpha, \delta, \beta, \bar{t})$ satisfy a (SNE) condition. Given $\mu, \nu > 0$, we will show how to find from $(v, \lambda, \alpha, \delta, \beta, \bar{t})$ a new tuple $(\tilde{v}, \tilde{\lambda}, \tilde{\alpha}, \tilde{\delta}, \tilde{\beta}, \tilde{t})$, with $\tilde{\beta}(s) = \nu s^\mu$ for all $s \geq 0$, still satisfying a (SNE) condition. It is sufficient to find a tuple $(\tilde{v}, \tilde{\lambda}, \tilde{\alpha}, \tilde{\delta}, \tilde{\beta}, \tilde{t})$ with $\tilde{\beta}(s) = s^\mu$ for all $s \geq 0$ and satisfying a (SNE) condition: the new tuple $(\nu\tilde{v}, \tilde{\lambda}, \nu\tilde{\alpha}, \nu\tilde{\delta}, \nu\tilde{\beta}, \tilde{t})$ will satisfy our claim. Since $\beta \in \mathcal{K}_\infty^1$ there exist $\beta_0 > 0$ for which $\beta_0 \leq \frac{s}{\beta(s)} \frac{d\beta}{ds}$ for all $s > 0$: this implies $b_0 s^{\beta_0} \leq \beta(s)$ for all $s \geq 1$ and for some $b_0 > 0$. For this reason and since by (P1) we can always augment $v(x, t)$ by a constant, we can assume $\beta(s) = \bar{\beta}(s) := b_0 s^{\bar{\beta}_0}$ for some $\bar{\beta}_0 \in (0, \mu)$ and for all $s > 0$. Indeed, for all $(x, t) \in \mathbb{R}^n \times [\bar{t}, +\infty)$

$$v(x, t) \geq \beta(\|x\|) \geq b_0 \|x\|^{\beta_0} - b_0 \geq b_0 \|x\|^{\bar{\beta}_0} - 2b_0$$

for any $\bar{\beta}_0 \in (0, \min\{\mu, \beta_0\})$. This implies that $(\bar{v}, \lambda, \alpha, \bar{\delta}, \bar{\beta}, \bar{t})$ satisfy a (SNE) condition with $\bar{v}(x, t) := v(x, t) + 2b_0$, $\bar{\beta}(s) := b_0 s^{\bar{\beta}_0}$, where $\bar{\beta}_0 \in (0, \mu)$, and $\bar{\delta}(s) := \delta(s) + 2\lambda b_0$, which proves the claim.

Let $h, a_0 > 0$ be such that $\frac{s}{\beta(s)^{\frac{1}{h}}} \leq (a_0 s)^{\frac{1}{h+1}}$ for all $\forall s > 0$, for instance $a_0 := b_0^{-(h+1)}$ and $h := \frac{\bar{\beta}_0}{\mu - \bar{\beta}_0}$. By considering the modified function

$$\tilde{v}(x, t) := (\bar{\beta}^{-1}(\bar{v}(x, t)))^\mu$$

and on account of the (PDI) condition on $(\bar{v}, \lambda, \alpha, \bar{\delta})$, for all $(x, \mathbf{d}, t) \in \mathbb{R}^n \times \mathcal{D} \times \mathbb{R}_{\geq}$ we have

$$\begin{aligned} & \frac{\partial \tilde{v}}{\partial x}(x, t) f(x, \mathbf{d}_t) + \frac{\partial \tilde{v}}{\partial t}(x, t) \leq \mu \tilde{v}^{\frac{\mu-1}{\mu}}(x, t) \frac{d\bar{\beta}^{-1}}{ds} \Big|_{v(x, t)} \times \\ & \times \left(-\lambda \bar{\beta}(\tilde{v}^{\frac{1}{\mu}}(x, t)) + \alpha(\|h(x, \mathbf{d}_t)\|) + \bar{\delta}(\|\mathbf{d}_t\|) \right) \\ & = \mu \left(\frac{s}{\bar{\beta}(s)} \frac{d\bar{\beta}}{ds} \Big|_{s=\tilde{v}^{\frac{1}{\mu}}(x, t)} \right)^{-1} \left[-\lambda \tilde{v}(x, t) + \frac{\tilde{v}(x, t)}{\bar{\beta}(\tilde{v}^{\frac{1}{\mu}}(x, t))} \times \right. \\ & \times \left. \left(\alpha(\|h(x, \mathbf{d}_t)\|) + \bar{\delta}(\|\mathbf{d}_t\|) \right) \right] \leq -\frac{\lambda \mu h}{(h+1)\bar{\beta}_0} \tilde{v}(x, t) \\ & + \frac{\mu(2a_0)^{\frac{1}{h}}}{\bar{\beta}_0 \lambda^{\frac{1}{h}}} \left(\alpha^{\frac{h+1}{h}}(\|h(x, \mathbf{d}_t)\|) + \bar{\delta}^{\frac{h+1}{h}}(\|\mathbf{d}_t\|) \right) \\ & := -\tilde{\lambda} \tilde{v}(x, t) + \tilde{\alpha}(\|h(x, \mathbf{d}_t)\|) + \tilde{\delta}(\|\mathbf{d}_t\|) \end{aligned} \quad (87)$$

where we used Young inequality and the inequality $|x+y|^p \leq 2^{p-1}(|x|^p + |y|^p)$ for all $x, y \in \mathbb{R}$ and $p \geq 1$. Moreover, on account of the (ULB) condition on $(\bar{v}, \bar{\beta}, \bar{t})$, we obtain $\tilde{v}(x, t) = (\bar{\beta}^{-1}(\bar{v}(x, t)))^\mu \geq (\bar{\beta}^{-1}(\bar{\beta}(\|x\|)))^\mu = \|x\|^\mu := \tilde{\beta}(\|x\|)$ for all $(x, t) \in \mathbb{R}^n \times [\bar{t}, +\infty)$. Therefore, $(\tilde{v}, \lambda, \tilde{\alpha}, \tilde{\delta}, \tilde{\beta}, \bar{t})$, with $\tilde{\beta}(s) = s^\mu$, satisfy a (SNE) condition which proves the claim. \triangleleft

B. Proof of (SP3)

By (P3) and given $\mu, \nu > 0$, from any tuple $(v, \lambda, \alpha, \delta, \beta, \bar{t})$ satisfying a (SNE) condition we can find a new tuple $(\nu \tilde{v}, \tilde{\lambda}, \nu \tilde{\alpha}, \nu \tilde{\delta}, \nu \tilde{\beta}, \bar{t})$, with $\tilde{\beta}(s) = s^\mu$ for all $s \geq 0$, satisfying a (SNE) condition. As it results from the proof of (P3), with $h := \frac{\bar{\beta}_0}{\mu - \bar{\beta}_0}$, $\psi := \frac{\mu(2a_0)^{\frac{1}{h}}}{\bar{\beta}_0 \lambda^{\frac{1}{h}}}$, $\bar{\beta}_0 \in (0, 1)$ and for some $b_0 > 0$ we have

$$\begin{aligned} \tilde{v}(x, t) &:= \left(\frac{1}{b_0} \right)^{\frac{\mu}{\bar{\beta}_0}} (v(x, t) + 2b_0)^{\frac{\mu}{\bar{\beta}_0}}, \\ \tilde{\alpha}(s) &:= \psi \alpha^{\frac{\mu}{\bar{\beta}_0}}(s), \quad \tilde{\delta}(s) := \psi \bar{\delta}^{\frac{\mu}{\bar{\beta}_0}}(s) \end{aligned}$$

Upon the (UUB) condition on $(v, \alpha, \zeta, \xi, \bar{t})$, it follows for all $t \geq \bar{t}$

$$\begin{aligned} & \frac{1}{\psi^{\frac{\bar{\beta}_0}{\mu}}} (\tilde{\alpha}(\|h(x, d)\|))^{\frac{\bar{\beta}_0}{\mu}} = \alpha(\|h(x, d)\|) \\ & \leq \zeta v(x, t) + \xi(\|d\|) \leq \zeta b_0 \tilde{v}^{\frac{\bar{\beta}_0}{\mu}}(x, t) + \xi(\|d\|) \end{aligned}$$

from which we obtain the new tuple $(\nu \tilde{v}, \tilde{\lambda}, \nu \tilde{\alpha}, \nu \tilde{\delta}, \nu \tilde{\beta}, \zeta, \xi, \bar{t})$ satisfying a (SSNE) condition with $\zeta > 0$ and $\xi \in \mathcal{K}_0$ defined as $\tilde{\zeta} := 2^{\frac{\mu}{\bar{\beta}_0}} \psi \zeta^{\frac{\mu}{\bar{\beta}_0}} b_0^{\frac{\mu}{\bar{\beta}_0}}$ and $\tilde{\xi}(s) := 2^{\frac{\mu}{\bar{\beta}_0}} \psi \xi^{\frac{\mu}{\bar{\beta}_0}}(s)$. \triangleleft

C. Proof of (44)

We prove the claim by transforming the tuple $(\Lambda^{-1}, \gamma, V, Z)$ into a new tuple $(\Lambda^{-1}, \tilde{\gamma}, \tilde{V}, \tilde{Z})$ satisfying a (HL)+(HD $_{\infty}$) condition with $\tilde{\gamma}_A$ and $\tilde{\gamma}_d$ satisfying (44) and (45), i.e. $\tilde{\gamma}_A(s) \geq 1$ and $\tilde{\gamma}_d(s) = s^{-\gamma_{d,0}}$ for all $s \geq 1$. Clearly, $\gamma_A(s) \geq \gamma_A(1)$ for all $s \geq 1$ and since $\gamma_d \in \mathcal{L}_{\infty}^1$ there exists $\gamma_{d,0} < 0$ such that $\gamma_d(s) \leq \gamma_d(1) s^{-\gamma_{d,0}}$ for all $s \geq 1$. Upon the form of the (HD $_{\infty}$) condition, we can increase γ_d and assume $\gamma_d(s) \equiv \gamma_d(1) s^{-\gamma_{d,0}}$ for all $s \geq 1$. Normalize (γ, V, Z) as $(\tilde{\gamma}, \tilde{V}, \tilde{Z})$ where $\tilde{\gamma}_A := \frac{\gamma_A}{\gamma_A(1)}$, $\tilde{\gamma}_d := \frac{\gamma_d}{\gamma_d(1)}$,

$\tilde{\Sigma}_A := \gamma_A^2(1) \Sigma_A$, $\tilde{\Sigma}_d := \gamma_A^2(1) \gamma_d^2(1) \Sigma_d$ and $\tilde{\Xi}_d := \gamma_d^2(1) \Xi_d$. The tuple $(\Lambda^{-1}, \tilde{\gamma}, \tilde{V}, \tilde{Z})$ satisfy a (HL)+(HD $_{\infty}$) condition with $\tilde{\gamma}_A$ and $\tilde{\gamma}_d$ as desired. \triangleleft

D. Proof of Theorem 6.1

We divide the proof in two steps. First, design a Lyapunov function for the estimation error system, finally we carry out the asymptotic convergence analysis.

1) *A Lyapunov function for the error system:* Let $\Delta e := e - \hat{e}$ be the estimation error and define a candidate Lyapunov function for the estimation error system as follows

$$\mathcal{L}(\Delta e, z) = \Delta e^\top \mathcal{P}(z) \Delta e \quad (88)$$

where $\mathcal{P}(z) = \Lambda^\top(z) \Pi \Lambda(z)$. In what follows, for simplicity, we write \mathbf{z}_t instead of $\hat{\mathbf{v}}_t + \hat{\gamma}$, Λ_t instead of $\Lambda(\hat{\mathbf{v}}_t + \hat{\gamma})$ and set $\sigma(\hat{e}, X, \Lambda) := \text{sat}_{e_{d\infty}}(\Lambda \hat{e}) + \text{sat}_{\hat{X}_{d\infty}}(\Lambda \hat{X})$. The estimation error system can be recast as

$$\begin{aligned} \dot{\Delta \mathbf{e}}_t &= A \Delta \mathbf{e}_t + \mathcal{E}_t^{(1)} + \mathcal{E}_t^{(2)} + \mathcal{E}_t^{(3)} \\ &\quad - \mathcal{H}(\mathbf{z}_t)(C \Delta \mathbf{e}_t + \mathcal{O}_t^{(1)} + \mathcal{O}_t^{(2)} + \mathcal{O}_t^{(3)}) \end{aligned} \quad (89)$$

with

$$\begin{aligned} \mathcal{E}_t^{(1)} &:= \mathcal{F}(\Lambda_t^{-1} \sigma(\Delta \mathbf{e}_t + \hat{\mathbf{e}}_t, \hat{\mathbf{X}}_t, \Lambda_t), 0) \\ &\quad - \mathcal{F}(\Lambda_t^{-1} \sigma(\hat{\mathbf{e}}_t, \hat{\mathbf{X}}_t, \Lambda_t), 0) \\ \mathcal{E}_t^{(2)} &:= \mathcal{F}(\Delta \mathbf{e}_t + \hat{\mathbf{e}}_t + \hat{\mathbf{X}}_t, \mathbf{d}_t) \\ &\quad - \mathcal{F}(\Lambda_t^{-1} \sigma(\Delta \mathbf{e}_t + \hat{\mathbf{e}}_t, \hat{\mathbf{X}}_t, \Lambda_t), \mathbf{d}_t) \\ \mathcal{O}_t^{(1)} &:= \mathcal{H}(\Lambda_t^{-1} \sigma(\Delta \mathbf{e}_t + \hat{\mathbf{e}}_t, \hat{\mathbf{X}}_t, \Lambda_t), 0) \\ &\quad - \mathcal{H}(\Lambda_t^{-1} \sigma(\hat{\mathbf{e}}_t, \hat{\mathbf{X}}_t, \Lambda_t), 0), \\ \mathcal{O}_t^{(2)} &:= \mathcal{H}(\Delta \mathbf{e}_t + \hat{\mathbf{e}}_t + \hat{\mathbf{X}}_t, \mathbf{d}_t) \\ &\quad - \mathcal{H}(\Lambda_t^{-1} \sigma(\Delta \mathbf{e}_t + \hat{\mathbf{e}}_t, \hat{\mathbf{X}}_t, \Lambda_t), \mathbf{d}_t) \end{aligned}$$

and

$$\begin{aligned} \mathcal{E}_t^{(3)} &:= \mathcal{F}(\Lambda_t^{-1} \sigma(\Delta \mathbf{e}_t + \hat{\mathbf{e}}_t, \hat{\mathbf{X}}_t, \Lambda_t), \mathbf{d}_t) \\ &\quad - \mathcal{F}(\Lambda_t^{-1} \sigma(\Delta \mathbf{e}_t + \hat{\mathbf{e}}_t, \hat{\mathbf{X}}_t, \Lambda_t), 0) \\ \mathcal{O}_t^{(3)} &:= \mathcal{H}(\Lambda_t^{-1} \sigma(\Delta \mathbf{e}_t + \hat{\mathbf{e}}_t, \hat{\mathbf{X}}_t, \Lambda_t), \mathbf{d}_t) \\ &\quad - \mathcal{H}(\Lambda_t^{-1} \sigma(\Delta \mathbf{e}_t + \hat{\mathbf{e}}_t, \hat{\mathbf{X}}_t, \Lambda_t), 0). \end{aligned}$$

Defining the homotopy $\text{Hom}(\vartheta, p_1, p_2) := \vartheta p_1 + (1 - \vartheta) p_2$, $\vartheta \in [0, 1]$ and $p_1, p_2 \in \mathbb{R}^n$, and

$$\begin{aligned} \mathcal{W}_t^{(1)} &:= \text{sat}_{e_{d\infty}}(\Lambda_t(\Delta \mathbf{e}_t + \hat{\mathbf{e}}_t)) + \text{sat}_{\hat{X}_{d\infty}}(\Lambda_t \hat{\mathbf{X}}_t), \\ \mathcal{W}_t^{(2)} &:= \text{sat}_{e_{d\infty}}(\Lambda_t(\hat{\mathbf{e}}_t)) + \text{sat}_{\hat{X}_{d\infty}}(\Lambda_t \hat{\mathbf{X}}_t), \quad \Delta \mathcal{W}_t := \mathcal{W}_t^{(1)} - \mathcal{W}_t^{(2)}, \\ \mathcal{V}_t^{(1)} &:= \Lambda_t(\Delta \mathbf{e}_t + \hat{\mathbf{e}}_t), \quad \mathcal{V}_t^{(2)} := \Lambda_t \hat{\mathbf{e}}_t, \quad \Delta \mathcal{V}_t := \mathcal{V}_t^{(1)} - \mathcal{V}_t^{(2)}, \end{aligned}$$

the functions $\mathcal{E}_t^{(1)}$, $\mathcal{O}_t^{(1)}$, $\mathcal{E}_t^{(3)}$ and $\mathcal{O}_t^{(3)}$ are easily decomposed as follows, according to the mean value theorem,

$$\begin{aligned} \mathcal{E}_t^{(1)} &= \left(\int_0^1 \frac{\partial \mathcal{F}(\Lambda_t^{-1} x, 0)}{\partial x} \Big|_{x=\text{Hom}(\vartheta, \mathcal{W}_t^{(1)}, \mathcal{W}_t^{(2)})} d\vartheta \right) \Delta \mathcal{W}_t \\ \mathcal{O}_t^{(1)} &= \left(\int_0^1 \frac{\partial \mathcal{H}(\Lambda_t^{-1} x, 0)}{\partial x} \Big|_{x=\text{Hom}(\vartheta, \mathcal{V}_t^{(1)}, \mathcal{V}_t^{(2)})} d\vartheta \right) \Delta \mathcal{V}_t \\ \Delta \mathcal{W}_t &= \left(\int_0^1 \frac{\partial \text{sat}_c(w)}{\partial w} \Big|_{w=\text{Hom}(\vartheta, \mathcal{V}_t^{(1)}, \mathcal{V}_t^{(2)})} d\vartheta \right) \Delta \mathcal{V}_t, \end{aligned} \quad (90)$$

$$\begin{aligned}\mathcal{C}_t^{(3)} &= \left(\int_0^1 \frac{\partial \mathcal{F}(\Lambda_t^{-1} \mathcal{W}_t^{(1)}, d)}{\partial d} \Big|_{d=\text{Hom}(\vartheta, \mathbf{d}_t, 0)} d\vartheta \right) \mathbf{d}_t \\ \mathcal{O}_t^{(3)} &= \left(\int_0^1 \frac{\partial \mathcal{H}(\Lambda_t^{-1} \mathcal{W}_t^{(1)}, d)}{\partial d} \Big|_{d=\text{Hom}(\vartheta, \mathbf{d}_t, 0)} d\vartheta \right) \mathbf{d}_t\end{aligned}$$

and by virtue of (B_∞) for some $\bar{t} \geq t_0$

$$\mathcal{C}^{(2)}(t) = 0, \mathcal{O}^{(2)}(t) = 0, \forall t \geq \bar{t}$$

(for simplicity set $\bar{t} = t_0$). On account of the first two inequalities in (HD_∞),

$$\begin{aligned}2\Delta \mathbf{e}_t^\top \mathcal{P}(\mathbf{z}_t) \mathcal{C}_t^{(1)}(t) &\leq \gamma_A(\mathbf{z}_t) \int_0^1 \left\{ \|\Pi \Lambda_t \Delta \mathbf{e}_t\|_{F_A^{-1}(x)}^2 + \right. \\ &\left. + \|\Lambda_t \Delta \mathbf{e}_t\|_{\Sigma_A(x)}^2 \right\}_{x=\text{Hom}(\vartheta, \mathcal{W}_t^{(1)}, \mathcal{W}_t^{(2)})} d\vartheta\end{aligned}$$

where we used $\sup_w \|\frac{\partial \text{sat}_c}{\partial w}(w)\| \leq 1$. Analogously,

$$\begin{aligned}2\Delta \mathbf{e}_t^\top \mathcal{P}(\mathbf{z}_t) \mathcal{O}_t^{(3)}(t) &\leq \gamma_A(\mathbf{z}_t) \int_0^1 \left\{ \|\Pi \Lambda_t \Delta \mathbf{e}_t\|_{F_d^{-1}(x,d)}^2 + \right. \\ &\left. \gamma_d^2(\mathbf{z}_t) \|\mathbf{d}_t\|_{\Sigma_d(x,d)}^2 \right\}_{(x,d)=(\mathcal{W}_t^{(1)}, \text{Hom}(\vartheta, \mathbf{d}_t, 0))} d\vartheta\end{aligned}$$

Similarly, on account of the last two inequalities in (HD_∞),

$$\begin{aligned}2\Delta \mathbf{e}_t^\top \mathcal{P}(\mathbf{z}_t) \mathcal{K}(\mathbf{z}_t) \mathcal{O}_t^{(1)} &\leq \gamma_A(\mathbf{z}_t) \int_0^1 \left\{ \|\Lambda_t \Delta \mathbf{e}_t\|_{\Xi_C(x)}^2 \right. \\ &\left. + \|R^{-1}(c) C \Lambda_t \Delta \mathbf{e}_t\|_{H_C^{-1}(x)}^2 \right\}_{x=\text{Hom}(\vartheta, \mathcal{W}_t^{(1)}, \mathcal{W}_t^{(2)})} d\vartheta\end{aligned}$$

and

$$\begin{aligned}2\Delta \mathbf{e}_t^\top \mathcal{P}(\mathbf{z}_t) \mathcal{K}(\mathbf{z}_t) \mathcal{O}_t^{(3)} &\leq \gamma_A(\mathbf{z}_t) \int_0^1 \left\{ \gamma_d^2(\mathbf{z}_t) \|\mathbf{d}_t\|_{\Xi_d(x,d)}^2 \right. \\ &\left. + \|R^{-1}(c) C \Lambda_t \Delta \mathbf{e}_t\|_{H_d^{-1}(x,d)}^2 \right\}_{(x,d)=(\mathcal{W}_t^{(1)}, \text{Hom}(\vartheta, \mathbf{d}_t, 0))} d\vartheta.\end{aligned}$$

By using (44), (SM) and the inequalities $\|\mathcal{W}_t^{(1)}\|, \|\mathcal{W}_t^{(2)}\| < n(e_{d_\infty} + \hat{X}_{d_\infty})$, $\|\text{Hom}(\vartheta, \mathcal{W}_t^{(1)}, \mathcal{W}_t^{(2)})\| < 2n(e_{d_\infty} + \hat{X}_{d_\infty})$ and $\|\text{Hom}(\vartheta, \mathbf{d}_t, 0)\| \leq d_\infty$ for all $t \geq 0$ and $\vartheta \in [0, 1]$, we obtain for $t \geq t_0$ and with $c := e_{d_\infty} + \hat{X}_{d_\infty}$

$$\overbrace{\mathcal{L}(\Delta \mathbf{e}_t, \mathbf{z}_t)} \leq \gamma_A(\mathbf{z}_t) \left\{ -\lambda_\pi \mathcal{L}(\Delta \mathbf{e}_t, \mathbf{z}_t) + \rho_\infty^2(c, d_\infty) \gamma_d^2(\mathbf{z}_t) \right\}. \quad (91)$$

2) *Error convergence analysis:* We will now prove the asymptotic bounds (58) and (60). First, since $\gamma_A \in \mathcal{K}_0$ we have $\lim_{t \rightarrow +\infty} \int_0^t \gamma_A(\mathbf{z}_s) ds = +\infty$, i.e. $\vartheta(t) := \int_0^t \gamma_A(\mathbf{z}_s) ds$ is a time scale. In what follows we use the overbars to denote any variable in the new time scale. Let $\overline{\mathcal{L}}_\vartheta$ denote $\mathcal{L}(\overline{\Delta \mathbf{e}}_\vartheta, \overline{\mathbf{z}}_\vartheta)$. From (91) we get in the time scale ϑ

$$\frac{d}{d\vartheta} \overline{\mathcal{L}}_\vartheta \leq -\lambda_\pi \overline{\mathcal{L}}_\vartheta + \rho_\infty^2(c, d_\infty) \gamma_d^2(\overline{\mathbf{z}}_\vartheta) \quad (92)$$

for $\vartheta \geq \vartheta_0 := \vartheta(t_0)$. Moreover,

$$\frac{d}{d\vartheta} \overline{\mathbf{v}}_\vartheta = \frac{1}{\gamma_A(\overline{\mathbf{z}}_\vartheta)} \left(-\lambda \overline{\mathbf{v}}_\vartheta + \alpha(\|h(\overline{\mathbf{x}}_\vartheta, \overline{\mathbf{d}}_\vartheta)\|) + \delta(d_\infty) \right) \quad (93)$$

and, on account of (31) and since $v(\mathbf{x}_t, t) \leq \hat{\mathbf{v}}_t + \hat{\gamma} = \mathbf{z}_t$ for $t \geq t_0$ by (18),

$$\alpha(\|h(\overline{\mathbf{x}}_\vartheta, \overline{\mathbf{d}}_\vartheta)\|) \leq \zeta \overline{\mathbf{z}}_\vartheta + \xi(d_\infty) \quad (94)$$

for $\vartheta \geq \vartheta_0$. From (45) and (92) we get for $\vartheta \geq \vartheta_0$

$$\frac{d}{d\vartheta} \overline{\mathcal{U}}_\vartheta \leq -q_d \overline{\mathcal{U}}_\vartheta, \quad \overline{\mathcal{U}}_\vartheta := \frac{\overline{\mathcal{L}}_\vartheta}{\gamma_d^2(\overline{\mathbf{z}}_\vartheta)} - \frac{2\rho_\infty^2(c, d_\infty)}{\lambda_\pi},$$

where $q_d := \lambda_\pi - 4|\gamma_{d,0}|(\zeta + \xi(d_\infty) + \delta(d_\infty)) > 0$ by (50). It follows from the definition of $\overline{\mathcal{L}}$ that for $\vartheta \geq \vartheta_0$

$$\frac{\|\overline{\Lambda}(\overline{\mathbf{z}}_\vartheta) \overline{\Delta \mathbf{e}}_\vartheta\|^2}{\gamma_d^2(\overline{\mathbf{z}}_\vartheta)} < \frac{2\rho_\infty^2(c, d_\infty)}{\lambda_\pi \lambda_{\min}^\Pi} + \frac{\overline{\mathcal{U}}_{\vartheta_0} + 1}{\lambda_{\min}^\Pi} e^{-q_d(\vartheta - \vartheta_0)}. \quad (95)$$

By applying $\limsup_{\vartheta \rightarrow +\infty}$ to both members of the above inequality and going back to the original time scale, we get the asymptotic bound (58).

Let's prove the asymptotic bound (60) under the additional condition (59). From (50) and (92) we get for $\vartheta \geq \vartheta_0$

$$\frac{d}{d\vartheta} \overline{\mathcal{F}}_\vartheta \leq -\lambda_\pi \overline{\mathcal{F}}_\vartheta, \quad \overline{\mathcal{F}}_\vartheta := \overline{\mathcal{L}}_\vartheta - \frac{2\rho_\infty^2(c, d_\infty) \gamma_d^2(\overline{\mathbf{z}}_\vartheta)}{\lambda_\pi},$$

so that for $\vartheta \geq \vartheta_0$

$$\overline{\mathcal{L}}_\vartheta < \frac{2\rho_\infty^2(c, d_\infty) \gamma_d^2(\overline{\mathbf{z}}_\vartheta)}{\lambda_\pi} + (\overline{\mathcal{L}}_{\vartheta_0} + 1) e^{-\lambda_\pi(\vartheta - \vartheta_0)}.$$

But, by the definition of $\overline{\mathcal{L}}$ and ⁶ since $\lambda_{\min}^{\Lambda(\overline{\mathbf{z}}_\vartheta)^\top \Lambda(\overline{\mathbf{z}}_\vartheta)} \geq \overline{\mathbf{z}}_\vartheta^{-2|\lambda_{\min}^W|}$ for $\vartheta \geq \vartheta_0$, it follows that $\|\overline{\Delta \mathbf{e}}_\vartheta\|^2 < \frac{\overline{\mathcal{L}}_\vartheta \overline{\mathbf{z}}_\vartheta^{2|\lambda_{\min}^W|}}{\lambda_{\min}^\Pi}$ for $\vartheta \geq \vartheta_0$. Moreover, on account of (44), (93) and (94)

$$\overline{\mathbf{z}}_\vartheta \leq e^{(\zeta + \xi(d_\infty) + \delta(d_\infty))(\vartheta - \vartheta_0)} \overline{\mathbf{z}}_{\vartheta_0} \quad (96)$$

for $\vartheta \geq \vartheta_0$. Since $\gamma_\infty := \sqrt{\sup_{s \geq 1} \gamma_d^2(s) s^{2|\lambda_{\min}^W|}} < +\infty$ by (45) and (59), we finally get for $\vartheta \geq \vartheta_0$

$$\|\overline{\Delta \mathbf{e}}_\vartheta\|^2 < \frac{2\rho_\infty^2(c, d_\infty) \gamma_\infty^2}{\lambda_\pi \lambda_{\min}^\Pi} + \frac{\overline{\mathbf{z}}_{\vartheta_0}^{2|\lambda_{\min}^W|} (\overline{\mathcal{L}}_{\vartheta_0} + 1)}{\lambda_{\min}^\Pi} e^{-r_d(\vartheta - \vartheta_0)}$$

where $r_d := \lambda_\pi - 2|\lambda_{\min}^W|(\zeta + \xi(d_\infty) + \delta(d_\infty)) > 0$ by (50). By applying $\limsup_{\vartheta \rightarrow +\infty}$ to both members of the above inequality and going back to the original time scale, we get the asymptotic bound (60). \square

E. Proof of Theorem 6.2

We claim that it is possible to transform the given tuple $(v, \lambda, \alpha, \delta, \beta, \zeta, \xi, \bar{t})$ satisfying a (SSNE) condition into a new tuple $(\tilde{v}, \tilde{\lambda}, \tilde{\alpha}, \tilde{\delta}, \tilde{\beta}, \tilde{\zeta}, \tilde{\xi}, \tilde{t})$ still satisfying (SSNE) and (68), which in particular implies

$$\limsup_{t \rightarrow +\infty} \|\Lambda(\hat{\mathbf{v}}_t + \hat{\gamma}) \mathbf{x}_t\| < c. \quad (97)$$

On account of this, with the other assumptions of Theorem 6.2 and the \mathcal{N} -filter and state norm estimate associated to the tuple $(\tilde{v}, \tilde{\lambda}, \tilde{\alpha}, \tilde{\delta}, \tilde{\beta}, \tilde{\zeta}, \tilde{\xi}, \tilde{t})$, the proof of Theorem 6.2 can be continued from this point exactly as the proof of Theorem 6.1.

We prove our claim. Let $\Delta_0 \in (0, 1)$. By (SP3) we can transform the tuple $(v, \lambda, \alpha, \delta, \beta, \zeta, \xi, \bar{t})$ into a new $(\tilde{v}, \tilde{\lambda}, \tilde{\alpha}, \tilde{\delta}, \tilde{\beta}, \tilde{\zeta}, \tilde{\xi}, \tilde{t})$ where $\tilde{\beta}$ is given in (67). This implies

⁶Since $W = T^\top W_D T$ with $W_D = \text{diag}\{\lambda_i \in \sigma(W), i = 1, \dots, n\}$ and for some orthonormal T and using standard properties of matrix exponentials, for all $s \geq 1$ it follows $\lambda_{\min}^{\Lambda^\top(s) \Lambda(s)} = \lambda_{\min}^{e^{2 \ln(s)} W_D} = \lambda_{\min}^{\text{diag}\{s^{2\lambda_i}, i=1, \dots, n\}} \geq s^{-2|\lambda_{\min}^W|}$.

for each $(x_0, \mathbf{d}) \in \mathbb{R}^n \times \mathcal{D}$ the existence of $t_0 \geq \bar{t}$ such that $\tilde{\beta}^{-1}(\hat{\mathbf{v}}_t + \hat{\gamma}) \geq \tilde{\beta}^{-1}(v(\mathbf{x}_t, t)) \geq \|\mathbf{x}_t\|$ for all $t \geq t_0$. On the other hand, since⁷ $\|\Lambda(s)\| \leq s^{-|\lambda_{max}^W|}$ for all $s \geq 1$, it follows $\|\Lambda((\hat{\mathbf{v}}_t + \hat{\gamma})^{\Delta})\mathbf{x}_t\| < c$ for all $t \geq t_0$ and $\bar{\Delta} \in (\Delta_0, 1]$. Applying the limsup to both parts of the last inequality, we get the claim. \triangleleft

REFERENCES

- [1] V. Andrieu, L. Praly, A. Astolfi, Homogeneous approximation, recursive observer design, and output feedback, *SIAM Journ. Contr. and Optim.*, vol. 47, no. 4, 1814-1850, 2008.
- [2] V. Andrieu, L. Praly, A. Astolfi, High gain observers with updated gain and homogeneous correction terms, *Automatica*, vol. 81, pp. 422-428, 2017.
- [3] D. Astolfi, R. Postoyan, D. Nesic, Uniting Observers, *IEEE Trans. Autom. Contr.*, DOI 10.1109/TAC.2019.2933395, 2019.
- [4] S. Battilotti, Robust observer design under measurement noise with gain adaptation and saturated estimates, *Automatica*, vol. 45, pp. 75-86, 2009.
- [5] G. Besancon, *Nonlinear observers and applications*, Springer, 2007.
- [6] N. Boizot, E. Busvelle, and J.-P. Gauthier, An adaptive high-gain observer for nonlinear systems, *Automatica*, vol. 46, no. 9, pp. 1483-1488, 2010.
- [7] S. Bonnabel, J.J. Slotine, A Contraction Theory-Based Analysis of the Stability of the Deterministic Extended Kalman Filter, *IEEE Trans. Autom. Contr.*, vol. 60, no. 2, pp. 565-569, 2015.
- [8] 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.
- [9] M. Shakarami, K. Esfandiari, A.A. Suratgar, H.A. Talebi, Peaking attenuation of high-gain observers usign adaptive technique, *IEEE Trans. Autom. Contr.*, vol. 65 (10), pp. 4215-4229, 2020.
- [10] J-P. Gauthier, I. Kupka, *Deterministic Observation Theory and Applications*, Cambridge Univ. Press, 2001.
- [11] H. K. Khalil, Cascade high-gain observers in output feedback control, *Automatica*, vol. 80, pp. 110-118, 2017.
- [12] 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.
- [13] M. Kawski, Families of Dilations and Asymptotic Stability, *Analysis of Control. Dyn. Syst.: Progress in Syst. and Contr. Theory*, Bonnard B., Bride B., Gauthier J.P., Kupka I. (eds.), vol 8. Birkhauser Boston, pp 285-294, 1991.
- [14] 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.
- [15] A. Luca, P. Rodriguez-Ayerbe, D. Dumur, Control of disturbed LPV systems in a LMI setting, *IFAC PapersOnLine*, vol. 44, no. 1, pp. 4149-4154, 2011.
- [16] G. Basile, G. Marro, On the observability of linear time-invariant systems with unknown inputs, *Journ. Opt. Theory and Appl.*, vol. 3, no. 6, pp. 410-415, 1969.
- [17] C. Qian, W. Lin, Output feedback control of a class of nonlinear systems: a nonseparation principle paradigm, *IEEE Trans. Autom. Contr.*, vol. 58, pp. 1710-1715, 2002.
- [18] A. Polyakov, Quadratic-like Stability of Nonlinear Homogeneous Systems, *IEEE 56th Conf. Dec. and Contr.*, vol. 81, pp. 577-582, December 12-15, MeULBourne, Australia, 2017.
- [19] 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.
- [20] K. Reif, F. Sonneman, and R. Unbehauen, An EKF-based nonlinear observer with a prescribed degree of stability, *Automatica*, vol. 34, pp. 1119-1123, 1998.
- [21] 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.
- [22] R. G. Sanfelice, L. Praly, Solution of a Riccati Equation for the Design of an Observer Contracting a Riemannian Distance, *54th IEEE Conf. Dec. and Contr.*, Osaka, Japan, pp. 4996-5001, 2015.
- [23] M.M. Seron, J.H. Braslavsky, G.C. Goodwin, *Fundamental Limitations in Filtering and Control*, Springer, 1997.
- [24] H. Shim, D. Liberzon, Nonlinear Observers Robust to Measurement Disturbances in an ISS Sense, *IEEE Trans. Autom. Contr.*, vol. 61, pp. 48-61, 2016.
- [25] Z. Wang, C. C. Lim, P. Shi, Y. Shen, $H_\infty L_\infty$ fault detection observer design for linear parameter-varying systems, *IFAC PapersOnLine*, vol. 50, no. 1, pp. 15271-15276, 2017.
- [26] A. Luca, P. Rodriguez-Ayerbe, D. Dumur, Control of disturbed LPV systems in a LMI setting, *IFAC PapersOnLine*, vol. 44, no. 1, pp. 4149-4154, 2011.
- [27] R. El Houda Tabeth, T. Raïssi, C. Combastel, A. Zolghadri, Design of interval observers for LPV systems subject to exogenous disturbances, *Europ. Contr. Conf.*, July 17-19, 2013, Zürich, Switzerland, pp. 1126-1131, 2013.



Stefano Battilotti was born in Rome, Italy, in 1962. In 1992 he joined the Dipartimento di Ingegneria Informatica, Automatica e Gestionale “Antonio Ruberti” (DIAG) of “Sapienza” in Rome (Italy), where since 2005 he is a Professor of Automatic Control. He authored more than 50 journal papers and the monograph “Noninteracting control with stability for nonlinear systems”, LNCIS Springer series. His research interests are currently focused on state estimation and control of nonlinear systems and networked systems.

⁷Since $W = T^T W_D T$ with $W_D = \text{diag}\{\lambda_i \in \sigma(W), i = 1, \dots, n\}$ and for some orthonormal T and using standard properties of matrix exponentials, for all $s \geq 1$ it follows $\|\Lambda(s)\| \leq \|e^{\ln(s)W_D}\| \leq s^{-|\lambda_{max}^W|}$.