



# Redesign of discrete-time nonlinear observers with state estimate constraints guaranteed in prescribed convex set

Daniele Astolfi, Pauline Bernard, Romain Postoyan, Lorenzo Marconi

## ► To cite this version:

Daniele Astolfi, Pauline Bernard, Romain Postoyan, Lorenzo Marconi. Redesign of discrete-time nonlinear observers with state estimate constraints guaranteed in prescribed convex set. 11th IFAC Symposium on Nonlinear Control Systems, NOLCOS 2019, Sep 2019, Vienna, Austria. hal-02411361

**HAL Id: hal-02411361**

**<https://hal.science/hal-02411361>**

Submitted on 14 Dec 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Redesign of discrete-time nonlinear observers with state estimate constraints guaranteed in prescribed convex set

Daniele Astolfi\* Pauline Bernard\*\* Romain Postoyan\*\*\*  
Lorenzo Marconi\*\*

\* *Univ Lyon, Université Claude Bernard Lyon 1, CNRS, LAGEPP  
UMR 5007, 43 boulevard du 11 novembre 1918, 69100 Villeurbanne,  
France.*

\*\* *CASY-DEI, University of Bologna, Viale Pepoli 3/2, 40123  
Bologna, Italy.*

\*\*\* *Université de Lorraine, CNRS, CRAN, 54516  
Vandoeuvre-lès-Nancy, France*

---

**Abstract:** We propose a technique to modify a given discrete-time (nonlinear) observer so that the state estimate remains in a given convex set, without altering the observer performances in terms of convergence and robustness to external disturbances. The proposed approach can be used to remove the peaking phenomenon or to attenuate the effect of impulsive outliers in the measures. It assumes that it is possible to execute a certain number of computations between any two sampling times in order to refine the current estimate and bring it back into the prescribed set. The proposed technique can be applied to any class of nonlinear observers for which a quadratic Lyapunov function is used to prove stability.

---

## 1. INTRODUCTION

In the context of observer design, it is well known that a wrong initialization of the estimate may induce very large estimation errors during transient behaviors. This phenomenon, also named *peaking*, see Khalil and Praly (2014), can be observed in both continuous and discrete time observers. For instance, peaking may be induced by large gains selected to guarantee fast convergence of the estimation error, or by dynamics possessing large overshoots caused by unstable poles on the open loop system, or by underdamping oscillatory behaviors, see, e.g., (Seron et al., 2012, §1, §8, §9). Moreover, large estimation errors can also occur in presence of large measurement disturbances of impulsive nature, denoted also as outliers, see Alessandri and Zaccarian (2018). This phenomena may make the observer impractical or unsafe to use during transient if not properly handled, see, e.g. Khalil and Praly (2014), and renders implementation tricky from a numerical perspective, in particular if the computational power is limited as in embedded systems, see, e.g., Henriksson (2006), Kopetz (2011). A simple solution could be to keep the observer dynamics intact but saturate its estimate. However, this approach does not solve the numerical issues and does not attenuate the effect of impulsive disturbances inducing undesired peaking in steady-state behavior. Another natural idea would be to saturate the observer dynamics and force the observer to remain in the prescribed set, but this may destroy the observer convergence properties as we show on an example.

For continuous-time systems, some solutions have been proposed, such as reducing the peaking in high-gain observers in Andrieu et al. (2016), Astolfi et al. (2018),

or ensuring the invariance of a convex set for the high-gain observers in Astolfi and Praly (2013); Maggiore and Passino (2003), and for linear Kalman filters with linear constraints in Berkane et al. (2018) and references therein. For discrete-time systems, many techniques have been developed to address the problem of outliers, as in Alessandri and Awawdeh (2016), Alessandri and Zaccarian (2018), De Palma and Indiveri (2017), Gandhi and Mili (2010). Yet, these approaches do not guarantee the state of the observer to remain in some given prescribed set.

We propose a technique to modify a given discrete-time observer that guarantees the estimate to remain in a prescribed convex compact set, while preserving the observer convergence properties. This technique is inspired from the convex design method first proposed in Astolfi and Praly (2013) for continuous-time observers and can be applied to any class of discrete-time nonlinear observer provided that a quadratic Lyapunov function for the error dynamics is known, see, for instance, Califano et al. (2003), Ciccarella et al. (1993), Ibrir (2007), Kazantzis and Kravaris (2001), Xie et al. (1994), Zemouche and Boutayeb (2013). We assume for this purpose that each discrete time corresponds to a sampling time of the plant and that some computations can be carried out between each consecutive sampling times. More precisely, if the estimate – following the initially given observer’s dynamics – exits the prescribed set at a sample time, we propose a refining strategy that brings it back into that viable set *before* the next sampling event. We guarantee that the Lyapunov function does not increase at these intermediate computations. Hence, the convergence properties of the observer are not altered.

The rest of the paper is organized as follows. The problem formulation is stated in Section 2 and the redesign approach presented in Section 3. Then, our main result is given in Section 4 and a numerical simulation is shown in Section 5. Finally, conclusions and future perspectives are discussed in Section 6.

*Notation:* We denote by  $\mathbb{R}$  the set of real numbers,  $\mathbb{R}_{\geq 0}$  is the set of non-negative real numbers,  $\mathbb{N}$  is the set of non-negative integers, and  $\mathbb{N}_{>0}$  is the set of positive integers. We define  $\mathbb{R}_{\geq 0}^n := \mathbb{R}_{\geq 0} \times \dots \times \mathbb{R}_{\geq 0}$  with  $n \in \mathbb{N}_{>0}$ . Given a set  $\mathcal{A} \subset \mathbb{R}^n$ ,  $\text{int}(\mathcal{A})$  stands for the interior of  $\mathcal{A}$ , namely the set of all points in  $\mathcal{A}$  that do not belong to its boundary. Given a real number  $x \in \mathbb{R}_{\geq 0}$ , we denote with  $\text{ceil}(x)$  the ceiling function, namely  $\text{ceil}(x) = \min\{y \in \mathbb{N} : y \geq x\}$ . We define  $\|\cdot\|$  as the standard Euclidean norm. Given a matrix  $A \in \mathbb{R}^{n \times n}$ , we denote by  $\|A\|$  its induced norm, and  $\text{eig}(A)$  the set of its eigenvalues. Given  $x \in \mathbb{R}$ , we define  $\text{sat}_r(x) := \max(\min(x, r), -r)$ . A continuous function  $\alpha : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  is of class  $\mathcal{K}$  if it is strictly increasing and  $\alpha(0) = 0$ .

## 2. PROBLEM STATEMENT

Consider the discrete-time system

$$x_{k+1} = f(x_k, u_k, d_k), \quad y_k = h(x_k, d_k) \quad (1)$$

where  $x_k \in \mathbb{R}^{n_x}$  is the state,  $u_k \in \mathbb{R}^{n_u}$  a known input,  $y_k \in \mathbb{R}^{n_y}$  is the output, and  $d_k \in \mathbb{R}^{n_d}$  is a perturbation acting on the dynamics and/or the measurement, at time  $k \in \mathbb{N}$ . The following assumption is stated.

*Assumption 1.* We are given compact subsets  $\mathcal{X}_0 \subset \mathbb{R}^{n_x}$ ,  $\mathcal{U} \subset \mathbb{R}^{n_u}$ ,  $\mathcal{D} \subset \mathbb{R}^{n_d}$ , and  $\mathcal{X} \subset \mathbb{R}^{n_x}$ , such that any trajectory of (1) initialized in  $\mathcal{X}_0$ , with input in  $\mathcal{U}$  and perturbation in  $\mathcal{D}$ , remains in  $\mathcal{X}$  for all  $k \in \mathbb{N}$ .

We suppose to know an observer of the form

$$\hat{x}_{k+1} = F(\hat{x}_k, u_k, y_k) \quad (2)$$

where  $\hat{x}_k \in \mathbb{R}^{n_x}$  is the state estimate at time  $k \in \mathbb{N}$ . By denoting the estimation error  $e := \hat{x} - x$  and the observer error dynamics as

$$\begin{aligned} e_{k+1} &= \Delta(x_k, e_k, u_k, d_k) \\ &:= F(e_k + x_k, u_k, h(x_k, d_k)) - f(x_k, u_k, d_k), \end{aligned} \quad (3)$$

the performances of observer (2) are characterized in the following assumption.

*Assumption 2.* There exist a matrix  $P = P^\top > 0$  and a map  $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}^{n_d} \rightarrow \mathbb{R}_{\geq 0}$ , such that, by defining

$$V(e) := e^\top P e \quad (4)$$

on  $\mathbb{R}^{n_x}$ , the following inequality holds for all  $(x, e, u, d) \in \mathcal{X} \times \mathbb{R}^{n_x} \times \mathcal{U} \times \mathcal{D}$

$$V(\Delta(x, e, u, d)) \leq \beta(V(e), d). \quad (5)$$

*Remark 1.* At this point, no particular assumption is made on the map  $\beta$  which characterizes the performances of the observer. For an exponentially stable observer which is input-to-state stable (ISS) with respect to the perturbation, the function  $\beta$  is typically of the form

$$\beta(v, d) = \rho v + \alpha(|d|) \quad (6)$$

with  $\rho \in [0, 1)$  and  $\alpha \in \mathcal{K}$ .  $\triangle$

In the following,  $\underline{\lambda}$  and  $\bar{\lambda}$  are the minimum and maximum eigenvalue of  $P$ , respectively, so that for all  $e$  in  $\mathbb{R}^{n_x}$ ,

$$\underline{\lambda}|e|^2 \leq V(e) \leq \bar{\lambda}|e|^2. \quad (7)$$

Although the plant state  $x_k$  is known to remain in  $\mathcal{X}$ , the estimate  $\hat{x}_k$  may leave  $\mathcal{X}$ , especially during transient behaviors. As a consequence, our goal is to modify the observer dynamics (2) to ensure the estimate  $\hat{x}_k$  remains in a given compact set  $\hat{\mathcal{X}} \supset \mathcal{X}$ , while preserving the original performances stated in Assumption 2.

In the following, we suppose that the discrete-time instances  $k$  correspond to sampling times  $t_k$  where a measurement  $y_k$  is available, and that between any two consecutive sampling events  $t_k$  and  $t_{k+1}$ , we are able to carry out  $\theta \in \mathbb{N}$  – to be defined – computations in order to refine the estimates given by (2). This assumption is realistic when system (1) is the *exact* or *approximate* discrete-time model obtained by sampling<sup>1</sup> a continuous-time system at instants  $t_k, t_{k+1}, \dots$ , if the digital controller has a computational frequency which is higher than the sampling frequency. In other words, we run the following algorithm depicted in Figure 1:

- (1) at instant  $t_k$ , we compute

$$\hat{x}_{k,1} = F(\hat{x}_k, u_k, y_k). \quad (8)$$

- (2) if  $\hat{x}_{k,1} \in \hat{\mathcal{X}}$ , we select  $\hat{x}_{k+1} = \hat{x}_{k,1}$ . Otherwise, if  $\hat{x}_{k,1} \notin \hat{\mathcal{X}}$ , we carry out  $\theta$  refining computational steps leading to  $\hat{x}_{k,i}$ ,  $i = 2 \dots \theta + 1$ , such that at the end of those steps  $\hat{x}_{k,\theta+1} \in \hat{\mathcal{X}}$ .

- (3) we finally define

$$\hat{x}_{k+1} = \hat{x}_{k,\theta+1}. \quad (9)$$

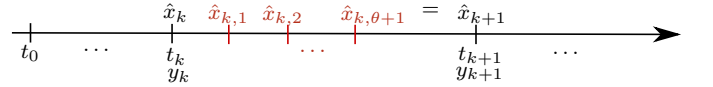


Figure 1. Discrete observer framework:  $\theta$  computational steps between each  $t_k$  and  $t_{k+1}$ .

We thus address the following problem.

*Problem 1.* For a given convex compact set  $\hat{\mathcal{X}} \supset \mathcal{X}$ , design a refining strategy, i.e. a map  $G : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$  and a number  $\theta \in \mathbb{N}$  so that, by combining (8)-(9) with

$$\hat{x}_{k,i+1} = G(\hat{x}_{k,i}), \quad i = 1 \dots \theta, \quad k \in \mathbb{N}, \quad (10)$$

the resulting new observer dynamics

$$\hat{x}_{k+1} = F_m(\hat{x}_k, u_k, y_k) \quad (11)$$

with  $F_m : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_x}$ , verify

- (1) the Lyapunov function (4) still satisfies (5) with  $\Delta$  replaced by

$$\begin{aligned} \Delta_m(x_k, e_k, u_k, d_k) \\ := F_m(e_k + x_k, u_k, h(x_k, d_k)) - f(x_k, u_k, d_k), \end{aligned} \quad (12)$$

- i.e. the performances of the observer are preserved,
- (2) for any initial conditions pair  $(x_0, \hat{x}_0) \in \mathcal{X}_0 \times \mathcal{X}$ , any input in  $\mathcal{U}$  and perturbation in  $\mathcal{D}$ , the corresponding solution to (11) lies in  $\hat{\mathcal{X}}$  for all  $k \in \mathbb{N}$ .

## 3. REDESIGN APPROACH

### 3.1 Convexity assumption

The method we propose requires convexity properties of the set  $\hat{\mathcal{X}}$  in which we want to keep the estimate trajectories. For this, let us define  $v_{\max}$  as

<sup>1</sup> See for instance Nešić et al. (1999) and references therein.

$$v_{\max} := \max_{(x, \hat{x}, d) \in \mathcal{X} \times \hat{\mathcal{X}} \times \mathcal{D}} \beta(V(\hat{x} - x), d), \quad (13)$$

with  $\beta$  given by Assumption 2, and the compact sets  $\Omega_{\max}$  and  $\hat{\mathcal{X}}_{\max}$  as

$$\Omega_{\max} := \{e \in \mathbb{R}^{n_x} : V(e) \leq v_{\max}\}, \quad (14)$$

$$\hat{\mathcal{X}}_{\max} := \{\hat{x} \in \mathbb{R}^{n_x} : \exists x \in \mathcal{X}, \hat{x} - x \in \Omega_{\max}\}. \quad (15)$$

According to Assumption 1, and in view of (5),  $v_{\max}$  defines the size of the largest Lyapunov level sets of  $V$  after a jump from  $x_k \in \mathcal{X}$ ,  $\hat{x}_k \in \hat{\mathcal{X}}$ , with  $u_k \in \mathcal{U}$  and  $d_k \in \mathcal{D}$ . In other words,  $\hat{x}_{k+1}$  is in  $\hat{\mathcal{X}}_{\max}$  and  $e_{k+1}$  is in  $\Omega_{\max}$  when  $(x_k, \hat{x}_k, u_k, d_k)$  is in  $\mathcal{X} \times \hat{\mathcal{X}} \times \mathcal{U} \times \mathcal{D}$  along the dynamics (2). We make the following assumption.

*Assumption 3.* There exist  $n_c \in \mathbb{N}$ , a  $C^1$  function  $c : \mathbb{R}^{n_x} \rightarrow \mathbb{R}_{\geq 0}^{n_c}$  and real numbers  $\varepsilon_{\max} > \varepsilon_{\min} > 0$  and  $\delta > 0$  such that, by letting  $c_i$  the  $i$ -th component of  $c$  and

$$\begin{aligned} \mathcal{C}_0 &:= \{x \in \mathbb{R}^{n_x} : c(x) = 0\}, \\ \mathcal{C}_{\min} &:= \{x \in \mathbb{R}^{n_x} : |c(x)| \leq \varepsilon_{\min}\}, \\ \mathcal{C}_{\max} &:= \{x \in \mathbb{R}^{n_x} : |c(x)| \leq \varepsilon_{\max}\}, \end{aligned} \quad (16)$$

the following properties hold.

- (1)  $\mathcal{X} \subset \mathcal{C}_0 \subset \mathcal{C}_{\min} \subset \hat{\mathcal{X}} \subset \mathcal{C}_{\max}$ .
- (2)  $\hat{\mathcal{X}}_{\max} \subseteq \mathcal{C}_{\max}$ .
- (3) Each function  $c_i : \mathbb{R}^{n_x} \rightarrow \mathbb{R}_{\geq 0}$  is convex on  $\mathcal{C}_{\max}$ , i.e.

$$c_i(\hat{x}) \leq c_i(x) + \frac{dc_i}{dx}(\hat{x})(\hat{x} - x) \quad \forall x, \hat{x} \in \mathcal{C}_{\max}. \quad (17)$$

$$(4) \sup_{x \in \mathcal{C}_{\max}} \left| \frac{dc}{dx}(x) \right| \leq \delta.$$

If the set  $\hat{\mathcal{X}}$  can be arbitrarily chosen, we have the following result.

*Lemma 1.* Consider  $Q = Q^\top > 0$  and  $\underline{r} > 0$ , such that  $\mathcal{X} \subset \{x \in \mathbb{R}^{n_x} : x^\top Q x \leq \underline{r}\}$ . Then, Assumption 3 is verified with  $\hat{\mathcal{X}} = \{x \in \mathbb{R}^{n_x} : x^\top Q x \leq \bar{r}\}$ ,

$$c(x) := \max \left\{ \frac{x^\top Q x}{\bar{r}} - 1, 0 \right\}, \quad (18)$$

for some  $\bar{r} > r > \underline{r}$  and  $\varepsilon_{\max} > \varepsilon_{\min} > 0$ .

If  $\mathcal{X}, \hat{\mathcal{X}}$  are imposed but a compact convex intersection of half-planes and quadrics fits between  $\mathcal{X}$  and  $\hat{\mathcal{X}}$ ,  $c$  can be taken multi-dimensional with each  $c_i$  built as in (18) from linear or quadratic maps, as showed in the next lemma.

*Lemma 2.* Suppose there exist real numbers  $\underline{r}_i$  and  $\bar{r}_i$   $i = 1 \dots n_x$  such that  $\mathcal{X} \subset \prod_{i=1}^{n_x} [\underline{r}_i, \bar{r}_i] \subset \hat{\mathcal{X}}$ . Then, Assumption 3 is verified with

$$\begin{aligned} c(x) &= (\underline{c}_1(x), \bar{c}_1(x), \dots, \underline{c}_{n_x}(x), \bar{c}_{n_x}(x)), \\ \underline{c}_i(x) &= \max \{ \underline{r}_i - x_i, 0 \}^2, \quad \bar{c}_i(x) = \max \{ x_i - \bar{r}_i, 0 \}^2, \end{aligned} \quad (19)$$

for some  $\varepsilon_{\max} > \varepsilon_{\min} > 0$ .

The proofs of Lemma 1 and 2 are omitted for space reasons.

### 3.2 Refining strategy

We introduce the map  $M : \mathcal{C}_{\max} \rightarrow \mathbb{R}^{n_x}$  defined as

$$M(\hat{x}) = -\gamma P^{-1} \frac{dc}{dx}(\hat{x})^\top c(\hat{x}), \quad (20)$$

with a positive scalar  $\gamma$ , and  $P$  coming from Assumption 2. As proved in Lemma 3 of Appendix A, if  $\gamma$  is sufficiently

small, modifying the estimate  $\hat{x}$  as  $\hat{x} + M(\hat{x})$  makes  $V(\hat{x} - x)$  decrease no matter where the plant state  $x$  actually is, with  $V$  the quadratic Lyapunov function defined in (4). Besides,  $V$  decreases strictly (by  $\gamma \varepsilon_{\min}^2$ ) when  $\hat{x}$  is outside  $\text{int}(\mathcal{C}_{\min})$ , with  $\mathcal{C}_{\min}$  defined in (16).

Since  $\mathcal{C}_{\min} \subseteq \hat{\mathcal{X}}$ , a possible strategy is therefore to repetitively use  $M$  to bring the estimate back in  $\hat{\mathcal{X}}$  at each iteration  $k$ . For that, consider  $v_{\min} > 0$  as

$$v_{\min} := \max \{ v \in \mathbb{R}_{>0} : \forall x \in \mathcal{X}, \{\hat{x} \in \mathbb{R}^{n_x} : V(\hat{x} - x) \leq v\} \subseteq \mathcal{C}_{\min} \}.$$

In other words,  $v_{\min}$  is the largest positive number such that the following property holds

$$x \in \mathcal{X} \text{ and } V(\hat{x} - x) \leq v_{\min} \implies \hat{x} \in \mathcal{C}_{\min} \subset \hat{\mathcal{X}}. \quad (21)$$

Since the true state  $x$  is known to be in  $\mathcal{X}$ , a way to bring  $\hat{x}$  back into  $\hat{\mathcal{X}}$  is thus to bring  $V$  below the threshold  $v_{\min}$ . This can be done by applying  $M$  at most  $\theta \in \mathbb{N}$  times, with  $\theta$  defined as

$$\theta := \text{ceil} \left( \frac{\max\{v_{\max} - v_{\min}, 0\}}{\gamma \varepsilon_{\min}^2} \right). \quad (22)$$

*Remark 2.* By definition of  $\hat{\mathcal{X}}_{\max}$ , it may occur that  $\hat{\mathcal{X}}_{\max} \subset \mathcal{C}_{\min}$  and that  $v_{\max} < v_{\min}$ . This means that after every jump of the original observer (2), the estimate remains in the desired set  $\hat{\mathcal{X}}$ . In such a case, no modification is needed, namely  $\theta = 0$  and all the analysis is trivially satisfied since no refining is needed.  $\triangle$

The redesign we propose consists in correcting the state estimate with the map  $M$  as long as the estimate is outside  $\mathcal{C}_{\min}$ , namely the function (10) is selected as

$$G(\hat{x}_{k,i}) = \begin{cases} \hat{x}_{k,i} + M(\hat{x}_{k,i}) & \text{if } \hat{x}_{k,i} \in \mathcal{C}_{\max} \setminus \mathcal{C}_{\min} \\ \hat{x}_{k,i} & \text{otherwise} \end{cases} \quad (23)$$

for  $i = 1 \dots \theta$ . As it can be noted by the expression of  $M$ , the correction (23) uses the gradient of the convex map  $c$ , namely  $-dc/dx$ , to bring  $\hat{x}_{k,i+1}$  closer to  $\hat{\mathcal{X}}$  along level sets of  $V$ . The recursive algorithm stops when we cross  $\mathcal{C}_{\min}$ . This strategy is depicted in Figure 2. Note that this could not be achieved in one iteration because  $\gamma$  needs to be sufficiently small to ensure that  $V$  decreases, as shown in Lemma 3 in the Appendix, which justifies the  $\theta$  steps.

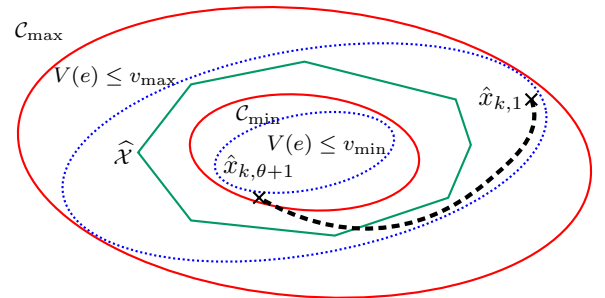


Figure 2. Refining strategy in (23). Red ellipsoids:  $\mathcal{C}_{\min}$ ,  $\mathcal{C}_{\max}$ . Green polygon:  $\hat{\mathcal{X}}$ . Dotted blue ellipsoids:  $V(e) \leq v_{\min}$  and  $V(e) \leq v_{\max}$ . Dashed black line: path of  $\hat{x}_{k,i+1}$ ,  $i = 0, \dots, \theta$ .

*Remark 3.* This redesign generates new observer dynamics (11) with a function  $F_m$  recursively defined through (23). If an explicit expression of  $F_m$  is available known, as in Berkane et al. (2018) for the particular case of linear dy-

namics with linear constraints, then the  $\theta$  computational steps of the refining algorithm are not necessary.  $\triangle$

#### 4. MAIN RESULT

With the refining strategy proposed in the previous section, we obtain the next result, addressing Problem 1.

*Theorem 1. Suppose Assumptions 2 and 3 hold and take*

$$0 < \gamma \leq \frac{\lambda}{\delta^2}. \quad (24)$$

*Then, the functions  $V$ ,  $F_m$ ,  $\Delta_m$ , defined by (4), (11), (12) with the algorithm (23), satisfy*

$$V(\Delta_m(x, e, u, d)) \leq \beta(V(e), d) \quad (25a)$$

$$F_m(\hat{x}, u, h(x, d)) \subset \hat{\mathcal{X}}, \quad (25b)$$

*for any  $(x, \hat{x}, u, d) \in \mathcal{X} \times \hat{\mathcal{X}} \times \mathcal{U} \times \mathcal{D}$  such that  $f(x, u, d) \in \mathcal{X}$ , and with  $e = \hat{x} - x$ .*

**Proof.** The proof of the theorem is omitted for space reasons. It can be deduced by direct application of Lemma 3 given in appendix. In particular, property (25a) can be proved by using the fact that at each step  $i \in \{1, \dots, \theta\}$ , the decrease of the Lyapunov function is preserved in view of inequality (A.1a). Property (25b) can be similarly proved by applying inequality (A.1b)  $\theta$  times and by using item 1) of Assumption 3.  $\square$

In Theorem 1, (25a) states that inequality (5) is preserved when using the modified dynamics (11). (25b) implies the invariance of  $\hat{\mathcal{X}}$  for the dynamics (11) at times  $k \in \mathbb{N}$ , namely that  $\hat{x}_{k+1} \in \hat{\mathcal{X}}$  if  $\hat{x}_k$  is computed according to the modified dynamics along (23), for any initial condition  $(x_k, \hat{x}_k) \in \mathcal{X} \times \hat{\mathcal{X}}$ , input in  $\mathcal{U}$  and perturbation in  $\mathcal{D}$ . We immediately deduce the next corollary.

*Corollary 1. Consider system (1) and suppose Assumptions 1, 2 and 3 hold with  $\beta$  of the type (6). There exist  $\nu > 0$  and a class  $\mathcal{K}$  function  $\alpha_0$  such that for any  $\gamma$  verifying (24) and for any solution  $x_k$  to (1) initialized in  $\mathcal{X}_0$ , any solution  $\hat{x}_k$  to (11) initialized in  $\hat{\mathcal{X}}$ , with input  $(u, d)$  in  $\mathcal{U} \times \mathcal{D}$  and  $y_k = h(x_k, d_k)$ , satisfies  $\hat{x}_k \in \hat{\mathcal{X}}$  and*

$$|\hat{x}_k - x_k| \leq \nu |\hat{x}_0 - x_0| \rho^{k/2} + \alpha_0 \left( \max_{i \leq k} |d_i| \right) \quad (26)$$

*for all  $k \in \mathbb{N}$ .*

The same ISS gain that we would have obtained without modifying the observer is ensured in (26), but we have gained the fact that  $\hat{x}$  remains in the prescribed set  $\hat{\mathcal{X}}$ .

In order to use the algorithm in (23) we need to be able to make (at most)  $\theta + 1$  recursive computations between  $t_k$  and  $t_{k+1}$ . As a consequence, depending on the sampling period  $t_{k+1} - t_k$  and the computational power of the system, this algorithm may have some limitations in practical applications. Note however that the Lyapunov function decreases even if we perform less than  $\theta + 1$  steps. Therefore, in case of limited computational power, it is still interesting to do as many  $\theta_m$  steps, with  $\theta_m \in \{0, \dots, \theta\}$ , as possible, as long as the estimate is outside of  $\mathcal{C}_{\min}$ : the Lyapunov function will decrease by  $\theta_m \gamma \varepsilon_{\min}^2$  which will make the algorithm converge faster and make the estimate stay closer to  $\hat{\mathcal{X}}$ .

The number  $\theta$  defined in (22) gives the largest number of refining steps needed to bring  $\hat{x}$  in  $\mathcal{C}_{\min}$ . However, this number may be significantly smaller, as the analysis is made on the conservative assumption that at each step, the decrease rate of  $V$  is  $\gamma \varepsilon_{\min}^2$ . But according to Lemma 3 in Appendix A, the true decrease is  $\gamma |c(\hat{x})|^2$ , which is larger than  $\gamma \varepsilon_{\min}^2$  when  $\hat{x}$  is far from  $\mathcal{C}_{\min}$ . In fact, the definition of the refining strategy (23) provides, from a computational point of view, a fast way to interrupt the algorithm once  $\hat{x}$  is in  $\mathcal{C}_{\min}$ .

An important observation is that although  $\hat{x}_k$  stays in  $\hat{\mathcal{X}}$ , the intermediary estimates  $\hat{x}_{k,i}$  can leave this set. However, according to the proof, they remain in  $\hat{\mathcal{X}}_{\max}$  defined in (15). Therefore,  $\hat{\mathcal{X}}_{\max}$  gives the magnitude of the numbers to be computed throughout the algorithm. This information can be useful to optimize the number encoding and increase numerical precision in case of limited memory, for instance in embedded systems. Actually, this set can be made more precise by considering the evolution of  $V$  given by (5) and the maximal initial error  $e_0 = \hat{x}_0 - x_0$ . For instance, in the case of Corollary 1,  $V$  remains smaller than

$$\tilde{v}_{\max} := V(e_{0,\max}) + \frac{1}{1-\rho} \alpha(d_{\max})$$

where  $e_{0,\max}$  and  $d_{\max}$  are bounds for the initial error and the disturbance respectively. Therefore,  $\hat{x}_{k,i}$  actually remain in

$$\hat{\mathcal{X}}_{\max} := \left\{ \hat{x} \in \mathbb{R}^{n_x} : \exists x \in \mathcal{X}, V(\hat{x} - x) \leq \min\{v_{\max}, \tilde{v}_{\max}\} \right\}$$

contained in  $\hat{\mathcal{X}}_{\max}$ . However, if the bounds  $e_{0,\max}$  and  $d_{\max}$  are small, it can happen that  $\tilde{v}_{\max} \leq v_{\max}$ , in which case  $\hat{x}_{k,i}$  are not guaranteed to be in a smaller set than if we had not modified the observer. Nevertheless, those bounds are extremely conservative, and it is very likely in practice that  $\hat{x}_{k,i}$  remain closer to  $\mathcal{X}$ .

All the aforementioned considerations are illustrated in Section 5 on an example.

*Remark 4.* The robustness property of the modified observer with respect to the perturbation  $d$  is characterized by (25a), and is therefore the same as with the initial observer. The added computational steps are not affected by perturbations, they are carried out independently from the plant and the observer. Therefore, the only robustness property we could be concerned about is the robustness of Theorem 1 with respect to numerical errors in the refining steps. The discrete dynamics (23) are not outer semicontinuous on the boundary of  $\mathcal{C}_{\min}$ , so that sequential compactness of solutions is not guaranteed. However, it turns out that the same result holds when allowing to use the correction  $M$  on the boundary, and we can show using the framework of discrete inclusion that the result of Theorem 1 is robust to computational errors.  $\triangle$

#### 5. SIMULATIONS

As a simple illustration, we consider the linear system

$$x_{k+1} = Ax_k, \quad y_k = Cx_k + d_k \quad (27)$$

with  $x_k \in \mathbb{R}^2$ , initial condition  $x_0 = (-1, -1)$ ,  $d_k \in \mathbb{R}$  a measurement disturbance and

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

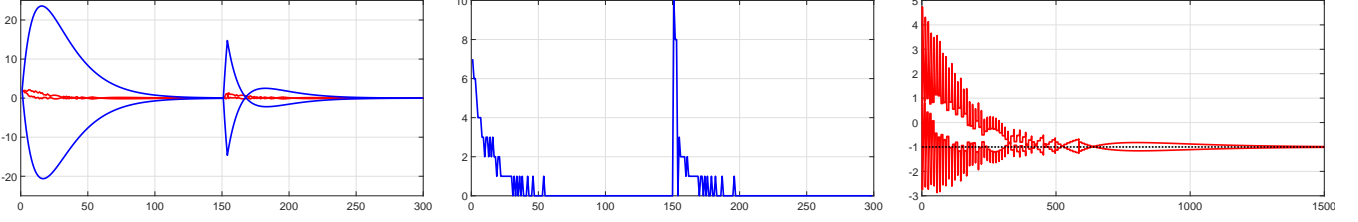


Figure 3. a) Error  $e_k$ , for  $k \in \{0, \dots, 300\}$ , provided by observer (28) without modification (blue line) and with modification (23) (red line) with  $\theta_m = 10$ . b) Number of refining steps of observer (28) with modification (23) at each  $k \in \{0, \dots, 300\}$ , with  $\theta_m = 10$ . c) Trajectory of  $\hat{x}_{k,i}$ , for  $k \in \{0, \dots, 150\}$  and  $i \in \{0, \dots, 10\}$ , of observer (28) with modification (23).

The corresponding solution to (27) is constant, namely  $x_k = (-1, -1)$  for all  $k \in \mathbb{N}$ , and lives therefore in the set  $\mathcal{X} = [-1, 1] \times [-1, 1]$ . Consider the observer

$$\hat{x}_{k+1} = A\hat{x}_k + L(y_k - C\hat{x}_k), \quad (28)$$

with  $L$  chosen such that the matrix  $A - LC$  is Schur, namely its eigenvalues lie in the open unit disk. The result of a simulation of observer (28) with  $\hat{x}_0 = (1, 1)$ ,  $L = [-1.87, 1874]$  ( $\text{eig}(A - LC) = \{0.95, 0.92\}$ ), and a disturbance  $d_k$  selected as

$$d_k = \begin{cases} 0, & 0 \leq k \leq 150 \text{ or } k \geq 154, \\ 3, & 150 < k < 154 \end{cases}$$

is given in Figure 3a). Two large peaking phenomena are present, caused by the wrong initial conditions and by the effect of the disturbance  $d$ .

We would like to preserve the convergence of observer (28) while keeping  $\hat{x}$  in the set  $\hat{\mathcal{X}} = [-r, r] \times [-r, r]$ , for some scalar  $r > 1$ . Note that we cannot implement the saturated observer

$$\hat{x}_{k+1} = \text{sat}_r(F(\hat{x}_k, u_k, y_k)), \quad (29)$$

where the function  $\text{sat}$  is taken component-wise, because the convergence is not preserved in absence of disturbances if  $r$  is taken too small. For instance, by selecting  $\hat{x}_0 = (1, 1)$  and  $r = 2$ , solutions to (29) remain stuck in  $\hat{x}_k = (2, -2)$  for all  $k > 0$ .

We thus apply the method presented in this paper. We know that for all  $\lambda > \max |\text{eig}(A - LC)|$ , there exists<sup>2</sup>  $P$  solution to

$$(A - LC)^\top P(A - LC) \leq \lambda^2 P, \quad P > 0.$$

We choose for instance  $\lambda = 0.955$ ,

$$P = 10^5 \cdot \begin{pmatrix} 1.259 & 1.319 \\ 1.319 & 1.381 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 0.039 & -0.037 \\ -0.037 & 0.036 \end{pmatrix}.$$

Then, Assumption 2 holds along the lines of Remark 1, so that we can follow Section 4 to modify the observer (28). By following Lemma 2, we choose the function  $c$  satisfying Assumption 3 as in (19) with

$$r_1 = -1.1, \quad \bar{r}_1 = 1.1, \quad r_2 = -1.1, \quad \bar{r}_2 = 1.1,$$

and  $\varepsilon_{\min} = 0.01$  which defines the set  $\mathcal{C}_{\min}$  that will be made invariant for the observer. As for  $\varepsilon_{\max}$  and  $\mathcal{C}_{\max}$ , notice that they are not used in the design apart from providing an upper bound on  $\gamma$  in (24) through  $\delta$  defined in Assumption 3. Here we take  $\gamma = 17$ . Instead of computing the theoretical number of steps  $\theta$  by using (22), we suppose that, due to computation limit constraints, we cannot carry out more than  $\theta_m$  computations in-between

<sup>2</sup> Because  $\frac{1}{\lambda}(A - LC)$  is Schur.

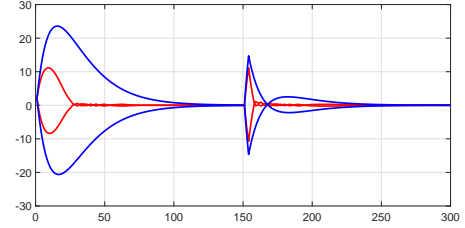


Figure 4. Error components  $e_k$ ,  $k \in \{0, \dots, 300\}$ , provided by observer (28) without modification (blue line) and with modification (23) (red line) with  $\theta_m = 3$ .

sample times to apply the refining algorithm (23). We performed two different scenarios:  $\theta_m = 3$  and  $\theta_m = 10$ . Figure 3 shows the behavior of the estimation error, the trajectory of the refining steps, and the number of refining steps required at each  $k$  for  $\theta_m = 10$  and for the same initial condition  $\hat{x}_0 = (1, 1)$ . The benefits of the refining algorithm are obvious: the convergence rate is fastened and the estimate is constrained in a set close to  $\mathcal{X}$  which is far smaller than the set where solutions to (28) naturally evolve. As a comparison, the results of the same simulation for  $\theta_m = 3$  are presented on Figure 4: the solution goes further from  $\mathcal{X}$  during the transients, but the benefits of the modification algorithm are still clearly visible. Table 1 shows the maximum values of each component in the three cases. Moreover, we see from Figures 3 and 4 how the modification algorithm enables to reject more efficiently the impulsive measurement disturbance  $d_k$ .

## 6. CONCLUSION

We proposed a technique that guarantees the estimate to remain in a prescribed convex set for a given observer's dynamics, without modifying its convergences properties. For this, we need to know a quadratic Lyapunov function for the error dynamics and we assume that, between any two samples, we are able to make a certain number of computations in order to refine the estimate of the observer. The proposed methodology allows one to attenuate the peaking phenomenon caused by wrong initial conditions and/or by measurement outliers, as illustrated by the numerical example. Future works aim at extending the proposed technique to observers with a state-dimension different from the plant's, as in e.g. Astolfi et al. (2018), observers with time-varying dynamics to handle Kalman-like designs as in Besançon (1999) or Boutayeb et al. (1997), and finally observers whose convergence is proved in other coordinates (Astolfi and Praly (2013), Bernard and Marconi (2018) Ciccarella et al. (1993)).



Observer without modification	
$\sup_k  \hat{x}_k^1  = 22.582$	$\sup_k  \hat{x}_k^2  = 21.611$
Observer with modification $\theta_m = 10$	
$\sup_k  \hat{x}_k^1  = 1.195$	$\sup_k  \hat{x}_k^2  = 1.238$
$\sup_{k,i}  \hat{x}_{k,i}^1  = 6.602$	$\sup_{k,i}  \hat{x}_{k,i}^2  = 4.783$
Observer with modification $\theta_m = 3$	
$\sup_k  \hat{x}_k^1  = 9.795$	$\sup_k  \hat{x}_k^2  = 8.366$
$\sup_{k,i}  \hat{x}_{k,i}^1  = 11.960$	$\sup_{k,i}  \hat{x}_{k,i}^2  = 10.358$

Table 1. Bounds on  $\hat{x}_k = (\hat{x}_k^1, \hat{x}_k^2)$  and refining steps  $\hat{x}_{k,i} = (\hat{x}_{k,i}^1, \hat{x}_{k,i}^2)$  for  $k \in \{0, \dots, 300\}$  and  $i \in \{0, \dots, \theta_m\}$ , with  $x_0 = (-1, -1)$ ,  $\hat{x}_0 = (1, 1)$ .

## Appendix A. PROPERTIES OF THE MAP $M$

*Lemma 3.* Suppose Assumptions 1 to 3 hold. If  $\gamma$  verifies (24), the functions  $V$  and  $M$  defined in (4) and (20) satisfy

$$V(e + M(\hat{x})) \leq V(e) \quad \forall (x, \hat{x}) \in \mathcal{X} \times \mathcal{C}_{\max} \quad (\text{A.1a})$$

$$V(e + M(\hat{x})) \leq V(e) - \gamma \varepsilon_{\min}^2 \quad \forall (x, \hat{x}) \in \mathcal{X} \times (\mathcal{C}_{\max} \setminus \text{int}(\mathcal{C}_{\min})) \quad (\text{A.1b})$$

where  $e = \hat{x} - x$ .

**Proof.** Take  $x$  in  $\mathcal{X}$  and  $\hat{x}$  in  $\mathcal{C}_{\max}$ . In view of (20),

$$V(e + M(\hat{x})) \leq V(e) - 2\gamma c(\hat{x})^\top \frac{dc}{dx}(\hat{x})(\hat{x} - x) + \gamma^2 c(\hat{x})^\top \frac{dc}{dx}(\hat{x}) P^{-1} \frac{dc}{dx}(\hat{x}) c(\hat{x}). \quad (\text{A.2})$$

By using the convexity properties in (17), we compute

$$\begin{aligned} -c(\hat{x})^\top \frac{dc}{dx}(\hat{x})(\hat{x} - x) &= -\sum_{i=1}^{n_c} c_i(\hat{x}) \frac{dc_i}{dx}(\hat{x})(\hat{x} - x) \\ &\leq -c(\hat{x})^\top (c(\hat{x}) - c(x)) \\ &\leq -c(\hat{x})^\top c(\hat{x}) \leq 0 \end{aligned} \quad (\text{A.3})$$

where we used the fact that  $c_i$  takes nonnegative values and  $c(x) = 0$  for all  $x \in \mathcal{X}$  by Assumption 3. By combining (A.2), (A.3), and by using (20), we obtain

$$\begin{aligned} V(e + M(\hat{x})) &\leq V(e) - \gamma c(\hat{x})^\top (2I - \gamma \frac{dc}{dx}(\hat{x}))^\top P^{-1} \frac{dc}{dx}(\hat{x}) c(\hat{x}) \\ &\leq V(e) - \gamma |c(\hat{x})|^2 \leq V(e) \end{aligned}$$

in which we used  $|P^{-1}| \leq \lambda^{-1}$ . Hence (A.1a) holds. Moreover, when  $\hat{x} \in \mathcal{C}_{\max} \setminus \text{int}(\mathcal{C}_{\min})$ , we have  $|c(\hat{x})| \geq \varepsilon_{\min}$  and hence (A.1b) holds.  $\square$

## REFERENCES

- Alessandri, A. and Awawdeh, M. (2016). Moving-horizon estimation with guaranteed robustness for discrete-time linear systems and measurements subject to outliers. *Automatica*, 67, 85–93.
- Alessandri, A. and Zaccarian, L. (2018). Stubborn state observers for linear time-invariant systems. *Automatica*, 88, 1–9.
- Andrieu, V., Priour, C., Tarbouriech, S., and Zaccarian, L. (2016). A hybrid scheme for reducing peaking in high-gain observers for a class of nonlinear systems. *Automatica*, 72, 138–146.
- Astolfi, D., Marconi, L., Praly, L., and Teel, A.R. (2018). Low-power peaking-free high-gain observers. *Automatica*, 98, 169–179.
- Astolfi, D. and Praly, L. (2013). Output feedback stabilization for siso nonlinear systems with an observer in the original coordinates. In *52nd IEEE Conference on Decision and Control*, 5927–5932.
- Berkane, S., Tayebi, A., and Teel, A.R. (2018). Hybrid constrained estimation for linear time-varying systems. *57th IEEE Conference on Decision and Control*.
- Bernard, P. and Marconi, L. (2018). Hybrid implementation of observers in plant’s coordinates with a finite number of approximate inversions and global convergence. *Submitted to Automatica*.
- Besançon, G. (1999). Further results on high gain observers for nonlinear systems. *38th IEEE Conference on Decision and Control*, 3, 2904–2909.
- Boutayeb, M., Rafaralahy, H., and Darouach, M. (1997). Convergence analysis of the extended kalman filter used as an observer for nonlinear deterministic discrete-time systems. *IEEE Transactions on Automatic Control*, 42(4), 581–586.
- Califano, C., Monaco, S., and Normand-Cyrot, D. (2003). On the observer design in discrete-time. *Systems & Control Letters*, 49(4), 255–265.
- Ciccarella, G., Dalla Mora, M., and Germani, A. (1993). Observers for discrete-time nonlinear systems. *Systems & Control Letters*, 20(5), 373–382.
- De Palma, D. and Indiveri, G. (2017). Output outlier robust state estimation. *International Journal of Adaptive Control and Signal Processing*, 31(4), 581–607.
- Gandhi, M.A. and Mili, L. (2010). Robust kalman filter based on a generalized maximum-likelihood-type estimator. *IEEE Transactions on Signal Processing*, 58(5), 2509–2520.
- Henriksson, D. (2006). *Resource-constrained embedded control and computing systems*. Ph.D. thesis, Department of Automatic Control, Lund University.
- Ibrir, S. (2007). Circle-criterion approach to discrete-time nonlinear observer design. *Automatica*, 43(8), 1432–1441.
- Kazantzis, N. and Kravaris, C. (2001). Discrete-time nonlinear observer design using functional equations. *Systems & Control Letters*, 42(2), 81–94.
- Khalil, H.K. and Praly, L. (2014). High-gain observers in nonlinear feedback control. *International Journal of Robust and Nonlinear Control*, 24(6), 993–1015.
- Kopetz, H. (2011). *Real-time systems: design principles for distributed embedded applications*. Springer Science & Business Media.
- Maggiore, M. and Passino, K.M. (2003). A separation principle for a class of non-ucio systems. *IEEE Transactions on Automatic Control*, 48(7), 1122–1133.
- Nešić, D., Teel, A.R., and Kokotović, P.V. (1999). Sufficient conditions for stabilization of sampled-data nonlinear systems via discrete-time approximations. *Systems & Control Letters*, 38(4-5), 259–270.
- Seron, M.M., Braslavsky, J.H., and Goodwin, G.C. (2012). *Fundamental limitations in filtering and control*. Springer Science & Business Media.
- Xie, L., Soh, Y.C., and De Souza, C.E. (1994). Robust kalman filtering for uncertain discrete-time systems. *IEEE Transactions on Automatic Control*, 39(6), 1310–1314.
- Zemouche, A. and Boutayeb, M. (2013). Comments on a note on observers for discrete-time lipschitz nonlinear systems. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 60(1).