

Alma Mater Studiorum Università di Bologna  
Archivio istituzionale della ricerca

Low-power peaking-free high-gain observers

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Astolfi, D., Marconi, L., Praly, L., Teel, A.R. (2018). Low-power peaking-free high-gain observers. AUTOMATICA, 98, 169-179 [10.1016/j.automatica.2018.09.009].

*Availability:*

This version is available at: <https://hdl.handle.net/11585/740933> since: 2020-02-28

*Published:*

DOI: <http://doi.org/10.1016/j.automatica.2018.09.009>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

*Daniele Astolfi, Lorenzo Marconi, Laurent Praly, Andrew R. Teel, **Low-power peaking-free high-gain observers**, Automatica, Volume 98, 2018, Pages 169-179, ISSN 0005-1098*

The final published version is available online at:

<https://doi.org/10.1016/j.automatica.2018.09.009>

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

*This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)*

***When citing, please refer to the published version.***

# Low-Power Peaking-Free High-Gain Observers

Daniele Astolfi <sup>a,b</sup>, Lorenzo Marconi <sup>a</sup>, Laurent Praly <sup>b</sup>, Andrew Teel <sup>c</sup>

<sup>a</sup>*CASY - DEI, University of Bologna, Italy*

<sup>b</sup>*MINES ParisTech, PSL Research University, CAS - Centre automatique et systèmes, Paris, France*

<sup>c</sup>*Electrical and Computer Engineering Department, University of California, Santa Barbara, CA 93106-9560 USA*

---

## Abstract

We propose a novel *peaking-free low-power high-gain observer* that may be used in place of standard high-gain observers. The proposed structure, with augmented dynamics, addresses the well-known problems of high-gain observers (numerical implementation difficulties and the peaking phenomenon) and improves the performance in the presence of high-frequency measurement noise.

*Key words:* High-gain observers, peaking, noise analysis

---

## 1 Introduction

High-gain observers appeared in the literature at the end of the 1980's and since then they have attracted a lot of research attention due to their simplicity and good performance in noise-free settings (see the survey [14] and references therein). See also their use in the separation principles [8], output feedback stabilization [27], output regulation [15] or fault detection [20].

In the design of a “standard” high-gain observer, the high-gain parameter, denoted as  $\ell$  throughout this paper, is usually powered up to  $n$ , with  $n$  denoting the dimension of the observed state. This fact raises numerical issues in the implementation when the state dimension is high or when the high-gain parameter has to be chosen large to achieve fast estimation. Furthermore, high-gain observers exhibit, during the transient, the so-called peaking phenomenon: namely the state of the observer suffers large peaks of a magnitude which is proportional to  $\ell^{n-1}$ . Last but not least, high-gain observers are known for having poor performance in the presence of high-frequency measurement noise, possibly making unusable in practice the state estimates especially when

the dimension  $n$  is very large. These problems are known from a very long time and many researchers have tried to overcome them, although without ever being able to solve them all together at the same time. Among the others, it is worth recalling [18], [23], [6] and [26], dealing with the peaking phenomenon and the works [1] [11], [19], [22] and [25] dealing with the problem of measurement noise.

A recent work [3] proposes a novel high-gain observer with a “low-power” implementation. In that paper, it is shown how to design a high-gain observer of dimension  $2n - 2$  for observable nonlinear systems with dimension  $n$ , which implements only gains proportional to  $\ell$  and  $\ell^2$  while preserving the same behaviours of a standard high-gain observer. This observer practically solves the aforementioned challenging problem of numerical implementation present in a standard high-gain observer. Moreover, it has been shown that the new observer structure substantially improves the sensitivity to high-frequency measurement noise. The proof of this fact has been presented in [3] only for linear systems, and shown by numerical simulation in the nonlinear case. The new low-power high-gain observer has been also shown to be effective for a much wider class of nonlinear systems, such as system possessing a non-strict feedback form (see [29]). It turns out that the new observer structure is effective in all those frameworks where standard high-gain observers are typically used, such as output feedback stabilization by nonlinear separation principle, [30], [29], and the output regulation [10]. Although the new observer structure solves the problem of numerical implementa-

---

\* Research supported in part by the European Project SHERPA (ICT 600958), by AFOSR grant FA9550-15-1-0155 and NSF grant ECCS-1508757.

*Email addresses:* [astolfi.daniele@gmail.com](mailto:astolfi.daniele@gmail.com) (Daniele Astolfi), [lorenzo.marconi@unibo.it](mailto:lorenzo.marconi@unibo.it) (Lorenzo Marconi), [Laurent.Praly@mines-paristech.fr](mailto:Laurent.Praly@mines-paristech.fr) (Laurent Praly), [teel@ece.ucsb.edu](mailto:teel@ece.ucsb.edu) (Andrew Teel).

tion, the peaking phenomenon is still present. This has been partially solved in [5]. In the latter, it has been shown that, by adding saturations at various levels in the observer structure, it is possible to remove the peaking from the first  $n - 1$  state estimates.

In this work, we combine the recent ideas of [3] and [5] to propose an observer of dimension  $2n - 1$  which is still “low power” (namely it uses only gains proportion  $\ell$  and  $\ell^2$ ) and yet eliminates the peaking phenomenon. Furthermore, we fully characterise the sensitivity to high-frequency measurement noise for nonlinear systems by showing the improvement with respect to standard observers. This is done by extending the analysis tool recently introduced in [4] in which the sensitivity to measurement noise has been characterised for standard high-gain observers. In this work, for the sake of simplicity, we focus on the same class of nonlinear systems in canonical observability form considered in [3], but similar results holds for the wider class of systems in feedback form [29].

The paper is organized as follows. After stating the framework and giving a brief introduction on the high-gain observer technique (Section 2), the main results are given in Section 3. A simulation example is given to show the performance of the observer in Section 4. All the proofs of the main results are given in the Appendix for the sake of clarity in the exposition.

**Notation.**  $\mathbb{R}$  denotes the field of real numbers and, for  $x \in \mathbb{R}^n$ ,  $|x|$  denotes the Euclidean norm of  $x$ . With  $s : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^m$  a bounded signal,  $\|s(\cdot)\|_{\infty} := \sup_{t \geq 0} |s(t)|$ . For  $i > 0$  we denote by  $A_i \in \mathbb{R}^{i \times i}$ ,  $B_i \in \mathbb{R}^{i \times 1}$ ,  $C_i \in \mathbb{R}^{1 \times i}$  a triplet in prime form, namely  $A_i$  is a shift matrix (all 1’s on the upper diagonal and all 0’s elsewhere),  $B_i = (0 \cdots 0 \ 1)$  and  $C_i = (1 \ 0 \cdots 0)$ . For  $r > 0$ , a saturation function  $\text{sat}_r : \mathbb{R} \rightarrow \mathbb{R}$  is any strictly increasing  $C^1$  function satisfying

$$\text{sat}_r(s) := s \quad \forall |s| \leq r, \quad |\text{sat}_r(s)| \leq r + 1 \quad \forall s \in \mathbb{R}.$$

With  $\mathcal{C}_{[0,1]}$  we denote the set of continuous functions  $s : \mathbb{R} \rightarrow [0, 1]$ .

## 2 The Framework and Highlights on High-Gain Observers

In this paper we deal with nonlinear single-input single-output systems that can be written, maybe after a change of coordinates, in the so-called *phase-variable form* (see [12])

$$\begin{aligned} \dot{x}_i &= x_{i+1}, & i &= 1, \dots, n-1, \\ \dot{x}_n &= \varphi(x, d(t)) \\ y &= x_1 + \nu(t) \end{aligned} \quad (1)$$

where the state  $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ , and  $t \mapsto d(t) \in \mathbb{R}^m$ ,  $m > 0$ , is any (unknown) bounded signal whose values range in a compact subset  $D \subset \mathbb{R}^m$  for all  $t \geq 0$ . In the following we suppose the state  $x(t)$  belongs to a given compact set  $X \subset \mathbb{R}^n$  for any  $t \geq 0$ . The function  $\varphi(\cdot, \cdot)$  is supposed to be locally Lipschitz in its first argument, namely we assume there exists a positive  $\bar{\varphi}_x$  such that

$$|\varphi(x_1, d) - \varphi(x_2, d)| \leq \bar{\varphi}_x |x_1 - x_2|$$

for all  $x_1, x_2 \in X$  and for all  $d \in D$ . In this framework  $d$  may represent uncertainties in the function  $\varphi(\cdot, \cdot)$  or unknown disturbances. All the forthcoming analysis can be extended, with the appropriate modifications, to the case in which the function  $\varphi(\cdot, \cdot)$  takes the form  $\varphi(x, d, t)$  where the dependence on  $t$  takes into account the effect of possible known inputs. For sake of simplicity, however, we do not consider this case.

The standard high-gain observer is given by

$$\begin{aligned} \dot{\hat{x}}_i &= \hat{x}_{i+1} + k_i \ell^i e_1, & i &= 1, \dots, n-1, \\ \dot{\hat{x}}_n &= \varphi_s(\hat{x}) + k_n \ell^n e_1, \end{aligned} \quad (2)$$

in which  $\hat{x} = (\hat{x}_1, \dots, \hat{x}_n)^T$  is the state,  $\ell$  is the high-gain parameter,  $e_1$  is the output injection term defined as

$$e_1 := y - \hat{x}_1,$$

$k_1, \dots, k_n$  are design coefficients and  $\varphi_s(\cdot)$  is any locally Lipschitz bounded function that agrees with  $\varphi(\cdot, 0)$  on a compact set  $X' \supset X$ , namely  $\varphi_s(x) = \varphi(x, 0)$  for all  $x \in X'$  and for all  $t \geq 0$ . The tuning of the observer involves choosing the design parameters  $k_i$ ’s so that, having defined the vector  $K := \text{col}(k_1, \dots, k_n)$ , the matrix  $A_n - KC_n$  is Hurwitz and by taking the high-gain parameter  $\ell$  large enough in relation to the Lipschitz constant of  $\varphi(\cdot, \cdot)$  on  $X \times D$ . In particular, it is possible to prove<sup>1</sup> that, by letting

$$\ell^* := 2 \bar{\varphi}_x |P|$$

in which  $P$  is the symmetric positive definite matrix solution of the Lyapunov equation

$$P(A_n - KC_n) + (A_n - KC_n)^T P = -I,$$

then for all  $\ell \geq \ell^*$  the error estimates provided by the

<sup>1</sup> This can be easily established by making the change of coordinates  $\chi = \text{col}(\chi_1, \dots, \chi_n)$ ,  $\chi_i = (\hat{x}_i - x_i)/\ell^{i-1}$  and by applying Lemma 3.

observer (2) satisfy the following bounds for all  $t \geq 0$ :

$$|\hat{x}_i(t) - x_i(t)| \leq c_1 \ell^{i-1} \exp(-c_2 \ell t) |\hat{x}(0) - x(0)| + \frac{c_3}{\ell^{n+1-i}} \|d(\cdot)\|_\infty + c_4 \ell^{i-1} \|\nu(\cdot)\|_\infty \quad (3)$$

$i = 1, \dots, n$ , for some positive constants  $c_i$ ,  $i = 1, \dots, 4$ , independent of  $\ell$ . As is clear from (3), one of the features of (2) is that the rate of convergence of the state estimate can be arbitrarily increased by augmenting the high-gain parameter  $\ell$  showing up in the exponential function, namely that the true value of the state variable can be practically recovered in an arbitrarily small amount of time. The term proportional to  $\ell^{i-1}$  multiplying the exponential, on the other hand, models the so-called peaking phenomenon governing the state estimate in the initial time instants. By this phenomenon the value of the estimation error assumes large values in the initial observation time if  $\ell$  is taken large. Hence, the smaller is the desired exponential decay, the larger is the peaking exhibited in the initial part of the transient. A further important feature of the observer (2) is that it is input-to-state stable (ISS) with respect to the disturbance inputs  $d$  and  $\nu$ . As for the disturbance  $d$ , in particular, the asymptotic gain on the  $i$ -th error variable is proportional to  $1/\ell^{n+1-i}$  and can be thus arbitrarily decreased by increasing the high-gain parameter. As for the measurement noise  $\nu$ , on the other hand, the asymptotic gain increases proportionally to  $\ell^{i-1}$ . The sensitivity to the class of bounded measurement noise signals, hence, tends to worsen with large values of the high-gain parameter with a polynomial term whose power increases with  $i$ . On top of everything, the main limit of the high-gain structure (2) is the presence of  $\ell$  powered up to the order  $n$ , which makes the numerical implementation of the observer a hard task for high-dimensional systems.

With reference to the sensitivity to measurement noise it is worth noting that the bound shown before refers to the so-called  $\mathcal{L}_\infty$  gain, namely characterise the sensitivity to the class of *bounded* disturbances. When considering the restricted class of *high-frequency* measurement noise the previous bound can be further refined by highlighting low-pass filtering properties of the observer (2). This high-frequency characterisation of the asymptotic gain has been fully characterised in [4] whose main result is briefly recalled hereafter.

We consider, in particular, the measurement noise as generated by the autonomous system

$$\begin{aligned} \varepsilon \dot{w} &= Sw, & w &\in \mathbb{R}^{n_w} \\ \nu &= Pw, \end{aligned} \quad (4)$$

where  $S \in \mathbb{R}^{n_w \times n_w}$  is a neutrally stable matrix,  $P$  is a row vector, and  $\varepsilon \in (0, 1)$  is a small parameter that parametrizes the frequency of the signal  $\nu(t)$ . System (4)

can be conveniently seen as generator of  $m > 0$  harmonics at frequencies  $\omega_i/\varepsilon > 0$ ,  $i = 1, \dots, m$ , namely, the matrices  $S$  and  $P$  take the form

$$S = \text{blkdiag}(S_1, \dots, S_m), \quad S_i = \begin{pmatrix} 0 & \omega_i \\ -\omega_i & 0 \end{pmatrix}$$

and  $P = ((0 \ 1) \ (0 \ 1) \ \dots \ (0 \ 1))$ . In the following we assume that  $w(0)$  ranges in a compact invariant set  $W$ . The main result proved in [4] is the following.

**Proposition 1** *Consider system (1)-(2) in the framework specified above with  $\nu$  generated by (4) and  $d(t) \equiv 0$ . Let  $\ell > 1$  be fixed so the bound (3) holds. Then, there exists a  $\varepsilon^*(\ell) > 0$  such that for all positive  $\varepsilon \leq \varepsilon^*(\ell)$  the following holds*

$$\lim_{t \rightarrow \infty} \sup |\hat{x}_i(t) - x_i(t)| \leq \varepsilon c \ell^i |w(0)| \quad i = 1, \dots, n.$$

with  $c$  a positive constant independent of  $\ell$ .

Once  $\ell$  is fixed the sensitivity of the estimation error to measurement noise thus decreases as  $\varepsilon$  takes smaller values, namely as higher frequency noise signals are considered, with an asymptotic gain proportional to  $\varepsilon$  for any  $i$ . For linear systems, this property immediately comes by frequency response arguments using the fact that the relative degree between the measurement noise  $\nu$  and the estimation error  $x - \hat{x}_i$  for (1)-(2) is unitary for all  $i = 1, \dots, n$ . The extension to nonlinear systems of the form (1) is more involved and can be found in [4].

### 3 Main Results

In this section we present the new high-gain observer structure preserving the main properties of the standard observer and solving the main problems discussed in the previous section. For sake of clarity the main results are presented in the next three subsections.

#### 3.1 Low-power high-gain observer

We start by presenting a high-gain observer of dimension  $2n - 1$  whose main feature is to have the high-gain parameter  $\ell$  that is powered just up to the order 2 regardless the value of  $n$ , thus overtaking one of the problems of the structure of (2). The observer structure strongly relies on the one presented in [3] that has dimension  $2n - 2$ . The motivation for extending the state of the observer by one with respect to the solution provided in [3] is to pave the way for the “peaking-free” solution that is then presented in Subsection 3.3.

The structure of the proposed observer is composed of  $n$  blocks, with each of the first  $n - 1$  blocks of dimension

2 and the last one of dimension 1. The two state components of the  $i$ -th block for  $i = 1, \dots, n-1$  are supposed to provide an estimate of  $(x_i, x_{i+1})$ , namely of the  $(i-1)$ -th and  $i$ -th time derivative of  $y$ , while the last block is meant to estimate the  $(n-1)$ -th time derivative of the output. The structure of the observer (see the next (5)) can be motivated as follows. If the  $i$ -th and  $(i+2)$ -th time derivative of  $y$ , i.e.  $x_i$  and  $x_{i+2}$ , were known, then the  $i$ -th block ( $i = 1, \dots, n-1$ ) could be implemented as a “nominal” high-gain observer for  $x_i$  and  $x_{i+1}$ , namely

$$\begin{aligned}\dot{\hat{x}}_i &= \eta_i + \ell \alpha_i (x_i - \hat{x}_i) \\ \dot{\eta}_i &= x_{i+2} + \ell^2 \beta_i (x_i - \hat{x}_i)\end{aligned}$$

where  $(\hat{x}_i, \eta_i)$  are estimates of  $(x_i, x_{i+1})$ ,  $\ell$  is the high gain parameter and  $(\alpha_i, \beta_i)$  are the observer parameters, with the entry  $x_{i+2}$  in the  $(n-1)$ -th block replaced by  $\varphi_s(x)$ . Similarly, the last 1-dimensional block could be implemented as

$$\dot{\hat{x}}_n = \varphi_s(x) + \ell \alpha_n (x_n - \hat{x}_n)$$

in which  $\alpha_n$  is a further design parameter and  $\hat{x}_n$  is meant to estimate  $x_n$ . Since  $x_i$  and  $x_{i+2}$  are not known, in fact, in the proposed observer their value is respectively replaced by  $\eta_{i-1}$  and  $\hat{x}_{i+1}$ , namely by the second and first component of the  $(i-1)$ -th and  $(i+1)$ -th block. By interconnecting the block observers this way, we get the proposed observer

$$\begin{aligned}\dot{\hat{x}}_i &= \eta_i + \alpha_i \ell e_i, & i = 1, \dots, n-1, \\ \dot{\hat{x}}_n &= \varphi_s(\hat{x}) + \alpha_n \ell e_n \\ \dot{\eta}_i &= \eta_{i+1} + \beta_i \ell^2 e_i, & i = 1, \dots, n-2, \\ \dot{\eta}_{n-1} &= \varphi_s(\hat{x}) + \beta_{n-1} \ell^2 e_{n-1}\end{aligned}\tag{5}$$

where  $\hat{x} = \text{col}(\hat{x}_1, \dots, \hat{x}_n) \in \mathbb{R}^n$ ,  $\eta = \text{col}(\eta_1, \dots, \eta_{n-1}) \in \mathbb{R}^{n-1}$  is the state,

$$\begin{aligned}\underline{\alpha} &:= \text{col}(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n \\ \underline{\beta} &:= \text{col}(\beta_1, \dots, \beta_{n-1}) \in \mathbb{R}^{n-1}\end{aligned}$$

are design parameters and  $\ell$  the high-gain parameter, and the variables  $e_i$ ,  $i = 1, \dots, n$  are defined as

$$\begin{aligned}e_1 &:= y - \hat{x}_1 \\ e_i &:= \eta_{i-1} - \hat{x}_i, & i = 2, \dots, n.\end{aligned}\tag{6}$$

The tuning of the design parameters  $\underline{\alpha}$  and  $\underline{\beta}$ , relies on a procedure that is different with respect to the one followed for the standard high-gain observers. In particular, having defined  $K_i := \text{col}(\alpha_i, \beta_i)$  and  $E_i := A_2 - K_i C_2$ ,

$i = 1, \dots, n-1$ , let the matrices  $M_i \in \mathbb{R}^{2i \times 2i}$ ,  $i = 1, \dots, n-1$ , and  $M_n \in \mathbb{R}^{(2n-1) \times (2n-1)}$  be recursively defined as  $M_1 := E_1$ ,

$$M_i := \begin{pmatrix} M_{i-1} & B_{2(i-1)} B_2^T \\ K_i B_{2(i-1)}^T & E_i \end{pmatrix} \quad i = 2, \dots, n-1$$

and

$$M_n := \begin{pmatrix} M_{n-1} & 0 \\ \alpha_n B_{2(n-1)}^T & -\alpha_n \end{pmatrix}.$$

With this notation in hand, the design parameters  $\underline{\alpha}$  and  $\underline{\beta}$  must be tuned in order to fulfil a “low-power stability requirement” that is formally defined in the following.

**Definition 1** (Low-power stability requirement). *We say that  $\underline{\alpha}$  and  $\underline{\beta}$  fulfil the “low-power stability requirement” if the resulting matrix  $M_n$  is Hurwitz, namely if there exists a  $P = P^T > 0$  such that*

$$PM_n + M_n^T P = -I. \tag{7}$$

It turns out that the eigenvalues of the matrix  $M_n$  can be arbitrary assigned by an appropriate choice of the design parameters; namely, the previous requirement can be always fulfilled (see Appendix A.1). With the matrix  $M_n$  Hurwitz and the high-gain parameter  $\ell$  taken sufficiently large, the estimation error  $\hat{x} - x$  provided by the observer (5)-(6) can be shown to fulfil the same bound (3) yielded by the standard observer. This is detailed in the next theorem in which we define  $\mathbf{x} \in \mathbb{R}^{2n-1}$  and  $\hat{\mathbf{x}} \in \mathbb{R}^{2n-1}$  as

$$\mathbf{x} := \text{col}((x_1, \dots, x_n), (x_2, \dots, x_n)), \quad \hat{\mathbf{x}} := \text{col}(\hat{x}, \eta).$$

**Theorem 1** *Consider the observer (5)-(6) and let the coefficients  $\underline{\alpha} \in \mathbb{R}^n$  and  $\underline{\beta} \in \mathbb{R}^{n-1}$  be chosen in order to fulfil the “low-power stability requirement”, with (7) fulfilled for some  $P = P^T > 0$ . Furthermore, let*

$$\ell^* := 2 \bar{\varphi}_x |P|.$$

*Then, there exist  $\mu_i > 0$ ,  $i = 1, \dots, 4$ , such that for any  $\ell > \ell^*$  the following bounds hold*

$$\begin{aligned}|\hat{x}_i(t) - x_i(t)| &\leq \ell^{i-1} \mu_1 \exp(-\ell \mu_2 t) |\hat{\mathbf{x}}(0) - \mathbf{x}(0)| + \\ &\quad \frac{\mu_3}{\ell^{n+1-i}} \|d(\cdot)\|_\infty + \mu_4 \ell^{i-1} \|\nu(\cdot)\|_\infty,\end{aligned}\tag{8}$$

for  $i = 1, \dots, n$ , and

$$\begin{aligned}|\eta_i(t) - x_{i+1}(t)| &\leq \ell^i \mu_1 \exp(-\ell \mu_2 t) |\hat{\mathbf{x}}(0) - \mathbf{x}(0)| + \\ &\quad \frac{\mu_3}{\ell^{n-i}} \|d(\cdot)\|_\infty + \mu_4 \ell^i \|\nu(\cdot)\|_\infty,\end{aligned}\tag{9}$$

for  $i = 1, \dots, n-1$ , for any initial condition  $(\hat{x}(0), \eta(0)) \in \mathbb{R}^n \times \mathbb{R}^{n-1}$  and for any  $x(t) \in \mathbb{R}^n$  and  $d(t) \in \mathbb{R}^m$  such that  $x(t) \in X$  and  $d(t) \in D$  for all  $t \geq 0$ .

The proof of this theorem is deferred to Appendix A.2. Note that the redundancy of the observer can be employed to obtain a double estimate of the state variables  $(x_2, \dots, x_n)$ , respectively given by  $(\hat{x}_2, \dots, \hat{x}_n)$  and  $\eta$ . Furthermore, we observe that the lower bound  $\ell^*$  of the high-gain parameter is formally equal to the one of the standard observer, namely it is proportional to the Lipschitz constant of  $\bar{\varphi}_x$  and to the norm of  $P$ . Regarding the latter, however, we observe that the fact that  $P$  is the solution of the Lyapunov equation associated to the matrix  $(A_n - KC_n) \in \mathbb{R}^{n \times n}$  for the standard observer and to  $M_n \in \mathbb{R}^{(2n-1) \times (2n-1)}$  for the new observer, the resulting value of  $\ell^*$  might be different. As clear from the bounds (8)-(9), the new observer preserves the same positive features of the standard observer in terms of an arbitrarily fast exponential decay rate of the estimation error and of an arbitrarily low asymptotic gain as far as the disturbance  $d$  is concerned, by overtaking the problem of (2) of having the high-gain parameter powered at  $n$ . On the other hand it does not eliminate the peaking phenomenon and it still has a sensitivity to the class of bounded measurement noise that depends on  $\ell$  polynomially in  $i$ .

### 3.2 Sensitivity to high-frequency noise

As at the end of Section 2, we now consider the measurement noise as generated by (4) and we characterise the asymptotic gain between  $\nu$  and the estimation error in terms of the parameter  $\varepsilon$ . The main objective is to show the benefit of the new observer in comparison with properties of the standard one as presented in Proposition 1. In this respect, the main feature of the observer (5)-(6) is that the relative degree between the “input”  $\nu$  and the  $i$ -th estimation error  $\hat{x}_i - x_i$  is one for  $i = 1$  (as for (2)) and then increases for higher values of  $i$ . More precisely, by defining  $m$  as

$$m := \left\lceil \frac{n+1}{2} \right\rceil \quad (10)$$

and considering a general case in which the function  $\varphi(x)$  is affected by  $x_1$ , it is easy to see that the relative degree in question is  $i$  for  $i = 1, \dots, m$  and  $n - i + 2$  for  $i = m+1, \dots, n$ . This property is at the basis of the next proposition whose proof is presented in Appendix A.3.

**Proposition 2** Consider system (1), (4)-(6) with  $d(t) \equiv 0$ . Let  $\underline{\alpha} \in \mathbb{R}^n$  and  $\underline{\beta} \in \mathbb{R}^{n-1}$  and  $\ell > 1$  be fixed according to the prescriptions of Theorem 1. Then, there exists a  $\varepsilon^*(\ell) > 0$  such that for all positive  $\varepsilon \leq \varepsilon^*(\ell)$  the

following holds

$$\begin{aligned} \lim_{t \rightarrow \infty} \sup |\hat{x}_i(t) - x_i(t)| &\leq \varepsilon^i c \ell^{2i-1} \|\nu(\cdot)\|_\infty \\ i &= 1, \dots, m, \\ \lim_{t \rightarrow \infty} \sup |\hat{x}_i(t) - x_i(t)| &\leq \varepsilon^{n-i+2} c \ell \|\nu(\cdot)\|_\infty \\ i &= m+1, \dots, n, \end{aligned}$$

for some constant  $c > 0$  independent of  $\ell$ .

The previous relation clearly shows that the new observer (as the standard one) behaves as “low-pass” filter, namely

$$\lim_{\varepsilon \rightarrow 0} \lim_{t \rightarrow \infty} \sup |\hat{x}_i(t) - x_i(t)| = 0.$$

In addition, the remarkable feature of the new observer is to have an asymptotic gain between the measurement noise and the  $i$ -th error component that is proportional to  $\varepsilon$  powered at a value that increases as long as “higher” components of the errors are considered, as opposed to the standard case in which the asymptotic gain depends on  $\varepsilon$  regardless the value of  $i$  (see Proposition 1). This fact, which is strongly related to the relative degree properties mentioned above, clearly shows that the new observer behaves better than the standard one as far as  $\varepsilon$  tends to zero, namely as far as high-frequency noise is concerned. The numerical analysis in Section 4 will provide further insight on the benefit of the new observer over the standard one.

### 3.3 Peaking-free low-power observer

In this subsection we finally show how the structure (5)-(6) can be modified in order to overtake also the problem of peaking while preserving the main features of the low-power observer presented before.

By bearing in mind the definition of saturation function and by defining  $r_i > 0$  as

$$r_i := \max_{x \in X} |x_i| \quad i = 1, \dots, n,$$

the low-power peaking-free observer takes the form (compare with (5)-(6))

$$\begin{aligned} \dot{\hat{x}}_i &= \eta_i + \alpha_i \ell e_i, & i &= 1, \dots, n-1, \\ \dot{\hat{x}}_n &= \varphi_s(\hat{x}) + \alpha_n \ell e_n, \\ \dot{\eta}_i &= \text{sat}_{r_{i+2}}(\eta_{i+1}) + \beta_i \ell^2 e_i, & i &= 1, \dots, n-2, \\ \dot{\eta}_{n-1} &= \varphi_s(\hat{x}) + \beta_{n-1} \ell^2 e_{n-1}, \end{aligned} \quad (11)$$

with

$$\begin{aligned} e_1 &:= y - \hat{x}_1 \\ e_i &:= \text{sat}_{r_i}(\eta_{i-1}) - \hat{x}_i, \quad i = 2, \dots, n, \end{aligned} \quad (12)$$

where  $\hat{x} = \text{col}(\hat{x}_1, \dots, \hat{x}_n) \in \mathbb{R}^n$ ,  $\eta = \text{col}(\eta_1, \dots, \eta_{n-1}) \in \mathbb{R}^{n-1}$  is the state,  $\underline{\alpha} := \text{col}(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$  and  $\underline{\beta} := \text{col}(\beta_1, \dots, \beta_{n-1}) \in \mathbb{R}^{n-1}$  are positive coefficients to be properly chosen, and  $\ell$  is the high-gain parameter.

The addition of saturation functions in (11) has the noteworthy effect of eliminating the peaking as clarified in the next proposition, but it imposes some restrictions on the choice of the design parameters  $\underline{\alpha}$  and  $\underline{\beta}$  with respect to the low-power stability requirement detailed in Definition 1. In particular, by bearing in mind the definitions of  $M_i$ ,  $K_i$ , and  $E_i$ , introduced in the previous subsection, let  $\Lambda_i : [0, 1] \rightarrow \mathbb{R}^{2i \times 2i}$ ,  $i = 1, \dots, n-1$ ,  $\Lambda_n \in \mathbb{R}^{2n-1 \times 2n-1}$  be continuous matrices defined as  $\Lambda_1 := M_1$ ,

$$\Lambda_i(s) := \begin{pmatrix} M_{i-1} & s B_{2(i-1)} B_2^T \\ K_i B_{2(i-1)}^T & E_i \end{pmatrix} \quad i = 2, \dots, n-1 \quad (13)$$

where  $s \in [0, 1]$ , and  $\Lambda_n := M_n$ . With this notations in mind, the design parameters  $\underline{\alpha}$  and  $\underline{\beta}$  must be tuned in order to fulfil a “low-power strong stability requirement” that is formally defined in the following.

**Definition 2** (Low-power strong stability requirement). *We say that  $\underline{\alpha}$  and  $\underline{\beta}$  fulfil the “low-power strong stability requirement” if the following holds:*

- $\alpha_i, i = 1, \dots, n$  and  $\beta_i, i = 1, \dots, n-1$ , are all positive;
- there exist  $P_i = P_i^T > 0$  and  $\mu_i > 0$  such that for all  $s \in [0, 1]$  the resulting  $\Lambda_i(s)$  fulfil

$$P_i \Lambda_i(s) + \Lambda_i(s)^T P_i \leq -\mu_i I \quad (14)$$

for all  $i = 1, \dots, n$ .

It turns out that the previous requirement can be always fulfilled by an appropriate choice of  $\underline{\alpha}$  and  $\underline{\beta}$ , with a constructive procedure that is presented in Appendix A.1.

**Proposition 3** *Consider the observer (11)-(12) with the design coefficients  $\underline{\alpha} \in \mathbb{R}^n$  and  $\underline{\beta} \in \mathbb{R}^{n-1}$  chosen so that the “low-power strong stability requirement” is fulfilled. Let  $(\hat{x}(0), \eta(0)) \in \hat{X} \times E$  with  $\hat{X} \times E$  an arbitrary compact set of  $\mathbb{R}^n \times \mathbb{R}^{n-1}$ . Then, the following holds:*

- (a) *there exist  $\bar{p}_i > 0$ ,  $i = 2, \dots, n$ , and, for each  $\bar{\nu} > 0$ ,*

*there exists a  $\bar{p}_1 > 0$  such that*

$$\begin{aligned} |\hat{x}_i(t) - x_i(t)| &\leq \bar{p}_i, \quad i = \dots, n \\ |\eta_i(t) - x_{i+1}(t)| &\leq \ell \bar{p}_i, \quad i = 1, \dots, n-1 \end{aligned} \quad (15)$$

*for all  $t \geq 0$ , for all  $\ell \geq 1$  and for all  $\nu(t)$  such that  $\|\nu(\cdot)\|_\infty \leq \bar{\nu}$ ;*

- (b) *there exist a  $\bar{\nu}$  and a  $\ell^* \geq 1$  such that for each  $\ell \geq \ell^*$  there exists a  $T > 0$  such that*

$$\|\nu(\cdot)\|_\infty \leq \frac{\bar{\nu}}{\ell^i} \Rightarrow \text{sat}_{r_{i+1}}(\eta_i(t)) = \eta_i(t)$$

*for all  $t \geq T$ ,  $i = 1, \dots, n-1$ .*

The observer (11)-(12), thus, guarantees that the estimation errors  $|\hat{x}_i(t) - x_i(t)|$ ,  $i = 1, \dots, n$ , do not peak with  $\ell$  (item (a) of the statement) and that the variables  $\eta_i$ ,  $i = 1, \dots, n-1$ , exit from saturation (item (b)) if the amplitude of the sensor noise is sufficiently small in relation to  $\ell^i$ . In particular, if  $\|\nu(\cdot)\|_\infty \leq \bar{\nu}/\ell^{n-1}$ , the observer (11)-(12) boils down, in finite time, to the low-power observer (5)-(6) by thus recovering all the asymptotic properties detailed in Theorem 1 and Proposition 2.

#### 4 A numerical example

We consider as example an uncertain harmonic oscillator described as

$$\ddot{\xi} + \omega^2 \xi = 0, \quad y = \xi + \nu(t)$$

where  $\omega$  is the unknown frequency and  $y$  is the measured output. We suppose that the output may be affected by high-frequency measurement noise  $\nu(t)$ . The purpose is to estimate the state  $\xi, \dot{\xi}$  and the unknown frequency  $\omega > 0$  from the output  $y$ . This system can be described in the state space by

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= -z_3 z_1 \\ \dot{z}_3 &= 0 \\ y &= z_1 + \nu(t) \end{aligned} \quad (16)$$

living in the open bounded subset of  $\mathbb{R}^3$

$$\mathcal{Z} = \left\{ (z_1, z_2, z_3) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} : d_1 < z_1^2 + z_2^2 < d_2, \right. \\ \left. d_3 < z_3 < d_4 \right\}$$

with  $d_1 < d_2$  and  $d_3 < d_4$  given positive numbers. In the state-space representation, (16)  $z_1$  coincides with  $\xi$ ,  $z_2$  with  $\dot{\xi}$ , and  $z_3$  with  $\omega^2$ . By following the results in [7], we

now immerse the third-order system (16) into a fourth-order system described in the phase-variable form (1). In particular, define

$$\begin{aligned}\varphi_0(z) &:= z_1, & \varphi_1(z) &:= z_2, \\ \varphi_2(z) &:= -z_1 z_3, & \varphi_3(z) &:= -z_2 z_3,\end{aligned}$$

and  $x = \Phi_4(z)$  where

$$\Phi_4(z) := \text{col}(\varphi_0, \varphi_1, \varphi_2, \varphi_3).$$

Simple computations show that the oscillator in  $x$  coordinates is described in the phase-variable form

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= x_4 \\ \dot{x}_4 &= \varphi(x) \\ y &= x_1 + \nu(t)\end{aligned}\tag{17}$$

where  $\varphi : \mathbb{R}^4 \rightarrow \mathbb{R}$  is defined as

$$\varphi(x) := -x_3 \frac{x_1 x_3 + x_2 x_4}{\max\{d_1, x_1^2 + x_2^2\}}.$$

It turns out that the function  $\Psi_4 : \Phi_4(\mathcal{Z}) \rightarrow \mathcal{Z}$  defined as

$$\Psi_4(\cdot) := \begin{pmatrix} \varphi_0 \\ \varphi_1 \\ -\frac{\varphi_2 \varphi_0 + \varphi_3 \varphi_1}{\varphi_0^2 + \varphi_1^2} \end{pmatrix}$$

is such that

$$z = \Psi_4(\Phi_4(z)) \quad \forall z \in \mathcal{Z}.$$

The observation of  $z$  can be thus accomplished by developing an observer for (17).

Numerical simulations have been run by considering  $d_1 = 0.5$ ,  $d_2 = 2$ ,  $d_3 = 0.5$ ,  $d_4 = 3$  so that

$$\begin{aligned}|x_1(t)| &< 4, & |x_2(t)| &< 5.5, \\ |x_3(t)| &< 12, & |x_4(t)| &< 16,\end{aligned}$$

for all  $t \geq 0$  and for any initial condition in the set  $\Phi_4(\mathcal{Z})$ . The observer (11)-(12) has been implemented with the coefficients  $\underline{\alpha}$  and  $\underline{\beta}$  chosen as as

$$\begin{aligned}\alpha_i &= 3 \quad i = 1, \dots, 4, \\ \beta_1 &= 6.4, \quad \beta_2 = 2.131, \quad \beta_3 = 0.7095\end{aligned}$$

so that the poles of the matrix  $M_n$  are  $-(1 + 0.2(i - 1))$  for  $i = 1, \dots, 4$ . The saturations level are fixed to  $r_1 = 5$ ,  $r_2 = 6$ ,  $r_3 = 14$ ,  $r_4 = 18$ ,  $r_5 = 55$ . We compared the observer (11)-(12) with a standard high-gain observer designed as (2) with  $k_1 = 5.99$ ,  $k_2 = 13.1778$ ,  $k_3 = 12.6034$ ,  $k_4 = 4.4156$  so that the roots of  $\lambda^4 + c_1 \lambda^3 + c_2 \lambda^2 + c_3 \lambda + c_4$  are in  $(-1, -1.33, -1.66, -2)$ .

In the simulation the initial conditions of the system (16) are  $(\xi(0), \dot{\xi}(0)) = (1, 0)$  and  $\omega = 1.58$  (namely  $z_3 = 2.5$ ) and the initial conditions of the two observers coincide with the origin. First, we simulated the nominal behaviours of the two observers in the absence of measurement noise, namely  $\nu(t) \equiv 0$ . The Table 1 shows the maximum peaking values of the state  $(\hat{x}, \eta)$  of the low-power peaking free high-gain observer (11)-(12) and the time needed to converge to an error sufficiently small, *i.e.* the time  $T_\epsilon$  such that

$$\sqrt{|\xi(t) - \hat{x}_1(t)|^2 + |\dot{\xi}(t) - \hat{x}_2(t)|^2} < \epsilon \quad \forall t \geq T_\epsilon,$$

for different values of  $\ell$ . Then, we simulated the behaviour of the two observers in presence of coloured measurement noise. The high-gain parameter is fixed to  $\ell = 15.5$  for the observer (11)-(12) and as  $\ell = 10$  for the (2) in order to practically match the convergence rate of the two observers. We repeated the simulations in two different scenarios. In the first we coloured a white-measurement noise with a band-pass filter (with band  $[10 - 50]Hz$ ) whereas in the second a high-pass filter has been used (with band  $[100 - \infty]Hz$ ). The Bode diagrams of the two filters are shown in Figure 1. Figures 2 and 3 show the transient behaviour of the low-power high-gain observer and the standard high-gain observer in the two scenarios. Despite the gain chosen for the standard high-gain observer (2) being smaller, it can be noticed a remarkable improvement of the sensitivity to high-frequency measurement noise of the new observer, in particular for the estimate of the frequency  $\hat{\omega}$  of the uncertain oscillator. The behaviours of the two observers in presence of medium-frequency measurement noise are substantially equivalent though the error on the frequency  $\hat{\omega}$  is slightly smaller.

## 5 Conclusion

In this paper a new class of nonlinear high-gain observers has been presented. Unlike classical high-gain observers, the proposed structure has the nice feature of having the high-gain parameter powered up to the order 2 regardless the dimension of the observed system, and it eliminates the so-called peaking phenomenon. Furthermore, superior performance in terms of sensitivity to high-frequency measurement noise has been shown. The

Low-power peaking-free high-gain observer				
	$\ell = 5$	$\ell = 10$	$\ell = 10^2$	$\ell = 10^3$
$T_{0.01}$	4.154	1.437	0.062	0.009
$\ \hat{x}_1(\cdot)\ _\infty$	1.46	1.46	1.48	1.57
$\ \hat{x}_2(\cdot)\ _\infty$	5.05	5.55	5.79	6.26
$\ \hat{x}_3(\cdot)\ _\infty$	6.52	6.52	6.41	6.49
$\ \hat{x}_4(\cdot)\ _\infty$	9.95	15.1	13.9	14.0
$\ \eta_1(\cdot)\ _\infty$	6.62	12.9	128	1308
$\ \eta_2(\cdot)\ _\infty$	14.0	27.0	443	5080
$\ \eta_3(\cdot)\ _\infty$	9.15	16.1	158	1727

Table 1

Maximum value of the state of the low-power peaking-free high-gain observer (11)-(12) when  $\nu(t) = 0$  for different values of  $\ell$ .

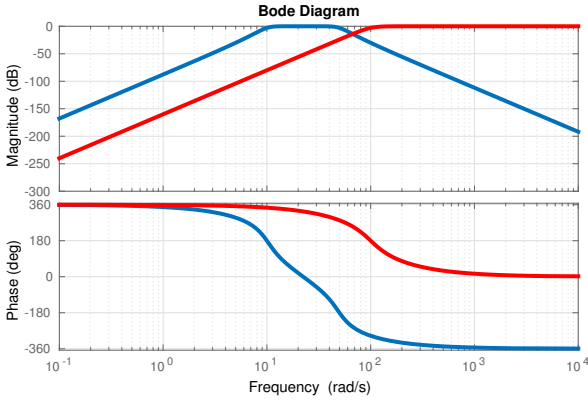


Fig. 1. Bode-digram of the medium-pass filter (blue line) and the high-pass filter (red line) used to colour the white-noise.

proposed structure can be used in place of the standard one in several settings where this class of observers is typically used, such as output feedback stabilization, output regulation, fault detection, and many others. In this paper we considered, for the sake of simplicity, observed systems in the so-called phase-variable form although the same ideas can be adopted to deal with more general observability forms.

## A Appendix

### A.1 Design of the coefficients fulfilling the “low-power (strong) stability requirement”

The tuning of the coefficients  $(\underline{\alpha}, \underline{\beta}) \in \mathbb{R}^n \times \mathbb{R}^{n-1}$  satisfying the “low-power stability requirement” (see Definition 1) can be done by means of a procedure that assigns the eigenvalues of the matrix  $M_n$ . This is presented in the next lemma that links to a constructive design procedure presented in [3].

**Lemma 1** *Let  $\mathcal{P}(\lambda) = \lambda^{2n-1} + m_1\lambda^{2n-2} + \dots + m_{2n-2}\lambda + m_{2n-1}$  be an arbitrary Hurwitz polynomial. There exists a choice of design coefficients  $(\underline{\alpha}, \underline{\beta}) \in \mathbb{R}^n \times \mathbb{R}^{n-1}$  such*

*that the characteristic polynomial of  $M_n$  coincides with  $\mathcal{P}(\lambda)$ .*

**Proof:** The triangular structure of the matrix  $M_n$  implies that  $\mathcal{P}(\lambda) = \mathcal{P}_{n-1}(\lambda)(\lambda - \alpha_n)$  with  $\mathcal{P}_{n-1}(\lambda)$  the characteristic polynomial of  $M_{n-1}$ . Using the constructive procedure in Lemma 1 of [3], it turns out that the coefficients  $(\alpha_1, \dots, \alpha_{n-1})$  and  $(\beta_1, \dots, \beta_{n-1})$  can be designed to assign an arbitrary polynomial  $\mathcal{P}_{n-1}(\lambda)$ . From this the result immediately follows.  $\square$

The MATLAB code for the design of the coefficients can be found in [2]. Alternatively, one can use the design procedure satisfying the “low-power strong stability requirement” which is presented below.

As for the design of the coefficients fulfilling the “low-power strong stability requirement” (see Definition 2), in the following we present a constructive procedure based on small-gain arguments that articulates in three steps.

**Step 1:** Choice of  $(\alpha_1, \beta_1)$ . The first step consists of designing  $(\alpha_1, \beta_1)$  to make the matrix  $\Lambda_1$  Hurwitz. Since  $\Lambda_1 = M_1 = E_1$ , the objective is trivially fulfilled by taking  $(\alpha_1, \beta_1)$  as any pair of positive numbers.

**Step  $i$ :** Choice of  $(\alpha_i, \beta_i)$ ,  $i = 2, \dots, n-1$ . We proceed by induction, by assuming that  $(\alpha_j, \beta_j)$ ,  $j = 1, \dots, i-1$ , have been fixed all positive so that  $M_{i-1}$  is Hurwitz and we show that there exists a choice of  $(\alpha_i, \beta_i)$  positive such that the resulting  $\Lambda_i(s)$  defined in (13) fulfils  $\Lambda_i(s)P_i + P_i\Lambda_i(s)^T \leq -\mu_i I$  for some  $P_i = P_i^T > 0$  and  $\mu_i > 0$  for all  $s \in [0, 1]$ , and  $M_i$  is Hurwitz. Towards this end, we regard the matrix  $\Lambda_i(s)$  as the state matrix of a system resulting from the feedback interconnection of a first Hurwitz system

$$\begin{aligned} \dot{x}_{i-1} &= M_{i-1}x_{i-1} + B_{2(i-1)}v_{i-1} \\ y_{i-1} &= B_{2(i-1)}^T x_{i-1} \end{aligned} \quad (\text{A.1})$$

with state  $x_i \in \mathbb{R}^{2i}$ , input  $v_{i-1} \in \mathbb{R}$  and output  $y_{i-1} \in \mathbb{R}$ , and a second subsystem

$$\begin{aligned} \dot{x}_i &= E_i x_i + K_i v_i \\ y_i &= B_2^T x_i \end{aligned}$$

with state  $x_i \in \mathbb{R}^2$ , input  $v_i \in \mathbb{R}$  and output  $y_i \in \mathbb{R}$ , under the feedback

$$v_{i-1} = s y_i, \quad v_i = y_{i-1}.$$

Since  $M_{i-1}$  is Hurwitz, there exists a  $\bar{P}_{i-1} = \bar{P}_{i-1}^T > 0$  such that the Lyapunov function  $\bar{V}_{i-1} = x_{i-1}^T \bar{P}_{i-1} x_{i-1}$  satisfies

$$\dot{\bar{V}}_{i-1} \leq -\epsilon_{i-1}|x_{i-1}|^2 + \varrho_{i-1}^2|v_{i-1}|^2 - |y_{i-1}|^2$$

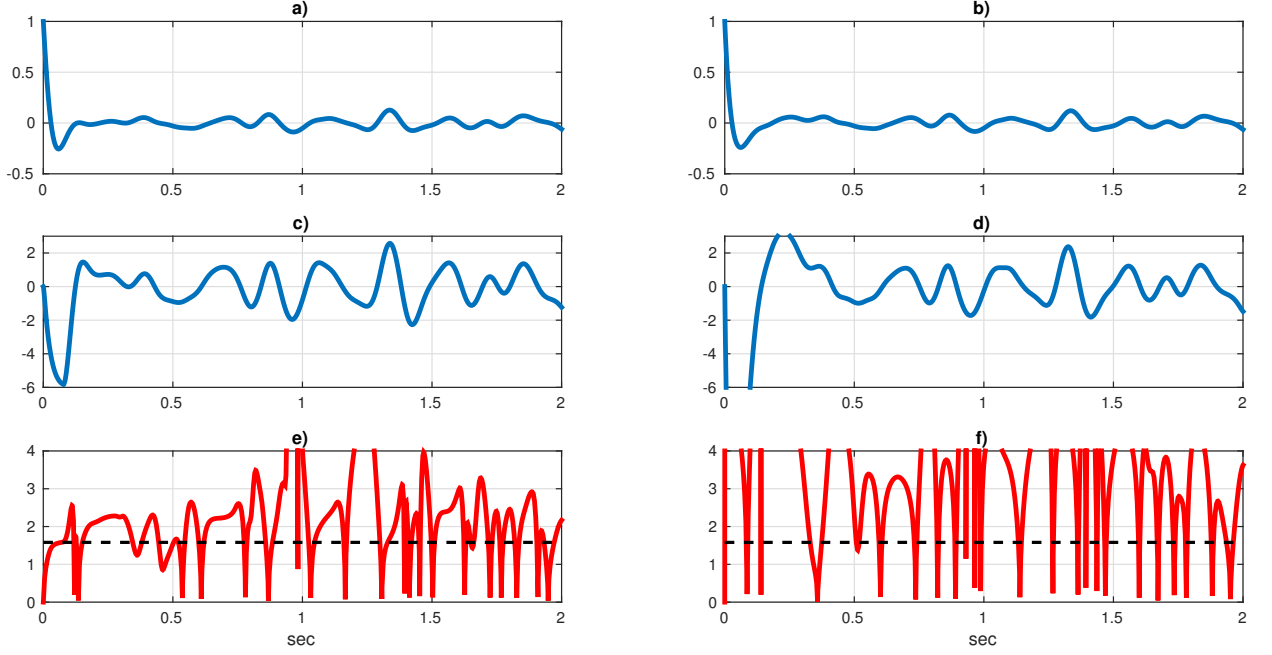


Fig. 2. Comparison between the behaviours of low-power peaking-free high-gain observer (11)-(12) with  $\ell = 15.5$  and the standard high-gain observer (2) with  $\ell = 10$  and white measurement noise coloured with a medium-pass filter. Plot **a**): behaviour of  $e_1(t) = \hat{x}_1(t) - \xi(t)$  for the low-power high-gain observer. Plot **b**): behaviour of  $e_1(t) = \hat{x}_1(t) - \xi(t)$  for the standard high-gain observer. Plot **c**): behaviour of  $e_2(t) = \hat{x}_2(t) - \dot{\xi}(t)$  for the low-power high-gain. Plot **d**): behaviour of  $e_2(t) = \hat{x}_2(t) - \dot{\xi}(t)$  for the standard high-gain. Plot **e**): behaviour of  $\hat{\omega}$  (red line) for the low-power high-gain observer and value of  $\omega$  (black line). Plot **f**): behaviour of  $\hat{\omega}$  (red line) for the standard high-gain observer. and value of  $\omega$  (black line).

for some positive  $\epsilon_{i-1}$  and  $\varrho_{i-1}$ . As for the system with state  $x_i$ , by bearing in mind the definitions of  $E_i$  and  $K_i$ , simple computations show that for all positive  $\alpha_i$  and  $\beta_i$  such a system is Hurwitz and its transfer function between input  $v_i$  and output  $y_i$  has  $\mathcal{H}_\infty$  gain equal to  $\beta_i/\alpha_i$ . Hence, by the bounded real lemma (see [17]), there exists  $\bar{P}_i = \bar{P}_i^T > 0$  such that the Lyapunov function  $\bar{V}_i = x_i^T \bar{P}_i x_i$  satisfies

$$\dot{\bar{V}}_i \leq -\epsilon_i |x_i|^2 + \gamma_i^2 |v_i|^2 - |y_i|^2$$

with  $\gamma_i := \beta_i/\alpha_i$ , for some positive  $\epsilon_i$ . Now, let

$$P_i := \text{blkdiag}(\bar{P}_{i-1}, c\bar{P}_i)$$

with  $c$  a positive constant yet to be defined, and consider the candidate Lyapunov function  $V_i = \xi_i^T P_i \xi_i$  in which  $\xi_i := \text{col}(x_{i-1}, x_i)$ . By bearing in mind that  $s \in [0, 1]$ , taking  $c \geq \varrho_{i-1}^2$  and designing  $(\alpha_i, \beta_i)$  positive so that  $\gamma_i^2 \leq 1/c$ , it turns out that

$$\begin{aligned} \dot{V}_i &\leq -\epsilon_{i-1} |x_{i-1}|^2 - c\epsilon_i |x_i|^2 + s\varrho_{i-1}^2 |y_i|^2 - |y_{i-1}|^2 \\ &\quad + c\gamma_i^2 |y_{i-1}|^2 - c|y_i|^2 \\ &\leq -\epsilon_{i-1} |x_{i-1}|^2 - c\epsilon_i |x_i|^2 \leq -\mu_i |\xi_i|^2 \end{aligned}$$

for some positive  $\mu_i$ . This, in particular, shows that (14) is fulfilled for all possible  $s \in [0, 1]$ . Furthermore, since  $M_i = \Lambda_i$  when  $s = 1$ , the previous arguments also show that  $M_i$  is Hurwitz.

**Step  $n$ :** choice of  $\alpha_n$ . We assume that  $(\alpha_j, \beta_j)$ ,  $j = 1, \dots, n-1$ , have been fixed all positive so that  $M_{n-1}$  is Hurwitz. Due to the fact that  $\Lambda_n = M_n$  and due to the triangular structure of the latter, the existence of a  $P_n = P_n^T > 0$  fulfilling (14) for  $i = n$  immediately follows by taking  $\alpha_n$  as any positive number.

In summary, the design of  $(\alpha, \beta)$  fulfilling the “low-power strong stability requirement” can be done by the following procedure:

- take  $(\alpha_1, \beta_1)$  as any pair of positive numbers;
- for all  $i = 2, \dots, n-1$ , compute recursively  $\alpha_i$  and  $\beta_i$  as any positive numbers such that  $\frac{\beta_i}{\alpha_i} \geq \frac{1}{\varrho_{i-1}}$  with  $\varrho_{i-1}$  the  $\mathcal{H}_\infty$  gain of system (A.1);
- take  $\alpha_n$  as any positive number.

#### A.2 Proof of Theorem 1

The proof follows the same idea of [3], with just minor adaptations due to the different dimension of the actual

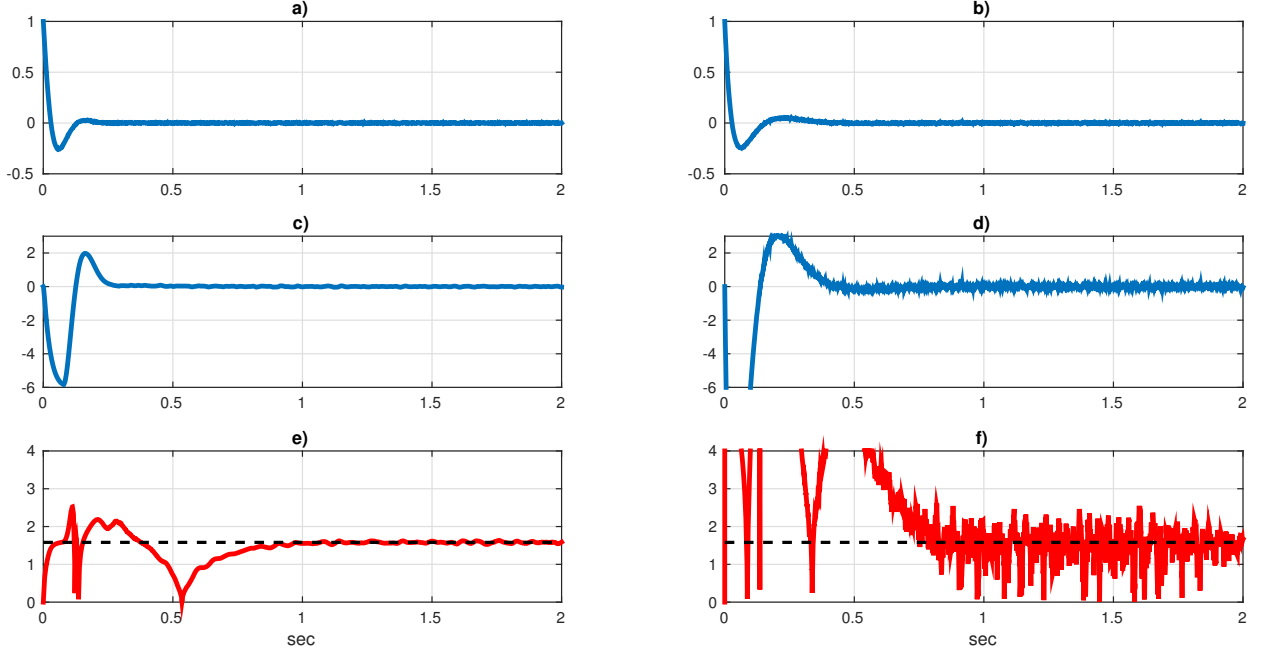


Fig. 3. Comparison between the behaviours of low-power peaking-free high-gain observer (11)-(12) with  $\ell = 15.5$  and the standard high-gain observer (2) with  $\ell = 10$  and white measurement noise coloured with a high-pass filter. Plot **a**): behaviour of  $e_1(t) = \hat{x}_1(t) - \xi(t)$  for the low-power high-gain observer. Plot **b**): behaviour of  $e_1(t) = \hat{x}_1(t) - \xi(t)$  for the standard high-gain observer. Plot **c**): behaviour of  $e_2(t) = \hat{x}_2(t) - \dot{\xi}(t)$  for the low-power high-gain. Plot **d**): behaviour of  $e_2(t) = \hat{x}_2(t) - \dot{\xi}(t)$  for the standard high-gain. Plot **e**): behaviour of  $\hat{\omega}$  (red line) for the low-power high-gain observer and value of  $\omega$  (black line). Plot **f**): behaviour of  $\hat{\omega}$  (red line) for the standard high-gain observer. and value of  $\omega$  (black line).

observer ( $2n-1$  instead of  $2n-2$ ), and therefore it is just sketched. Let  $\tilde{\chi}_i \in \mathbb{R}^2$  for  $i = 1, \dots, n-1$  and  $\tilde{\chi}_n \in \mathbb{R}$ , be defined as

$$\tilde{\chi}_i = \begin{pmatrix} \tilde{\chi}_{i1} \\ \tilde{\chi}_{i2} \end{pmatrix} := \begin{pmatrix} \ell^{-(i-1)}(\hat{x}_i - x_i) \\ \ell^{-i}(\eta_i - x_{i+1}) \end{pmatrix} \quad i = 1, \dots, n-1, \\ \tilde{\chi}_n := \ell^{-(n-1)}(\hat{x}_n - x_n). \quad (\text{A.2})$$

By letting  $\tilde{\chi} := \text{col}(\tilde{\chi}_1, \dots, \tilde{\chi}_n)$ , simple computations show that

$$\dot{\tilde{\chi}} = \ell M_n \tilde{\chi} + \frac{1}{\ell^{n-1}} \bar{B} \Delta_\varphi(\tilde{\chi}, x, d) + \ell \bar{K}_1 \nu(t)$$

where  $\bar{B} = \text{col}(B_{2n-2}, 1)$ ,  $\bar{K}_1 := \text{col}(K_1, 0, \dots, 0)$ , and

$$\Delta_\varphi(\tilde{\chi}, x, d) := \varphi_s(C_2 \tilde{\chi}_1 + x_1, \dots, \ell^{n-2} C_2 \tilde{\chi}_{n-1} + x_{n-1}, \ell^{n-1} \tilde{\chi}_n + x_n) - \varphi(x_1, \dots, x_n, d).$$

By bearing in mind that  $\varphi(\cdot, \cdot)$  is locally Lipschitz in  $x$  and that  $\varphi_s(\cdot)$  is locally Lipschitz, bounded and agrees

with  $\varphi(\cdot, 0)$  on  $X$ , it turns out that

$$\begin{aligned} |\varphi(x, d) - \varphi_s(\hat{x})| &= |\varphi(x, d) - \varphi(x, 0) + \varphi(x, 0) - \varphi_s(\hat{x})| \\ &\leq \bar{\varphi}_x |x - \hat{x}| + R \end{aligned}$$

for all  $(x, \hat{x}, d) \in X \times \mathbb{R}^n \times D$  and with  $R := 2 \max_{\{x \in X, d \in D\}} |\varphi(x, d)|$ . Furthermore, it is not hard to see that

$$|\Delta_\varphi(\tilde{\chi}, x, d)| \leq \ell^{n-1} \bar{\varphi}_x |\tilde{\chi}| + R.$$

Using now the fact that  $M_n$  is Hurwitz and by applying Lemma 3 in Appendix A.5, it is possible to claim that for all  $\ell \geq \ell^*$ , with  $\ell^*$  indicated in the statement of the theorem, the following bound holds

$$|\tilde{\chi}(t)| \leq \mu_1 \exp(-\ell \mu_2 t) |\tilde{\chi}(0)| + \frac{\mu_3}{\ell^n} \|d(\cdot)\|_\infty + \mu_4 \|\nu(\cdot)\|_\infty$$

for some positive  $\mu_i$ ,  $i = 1, \dots, 4$  independent of  $\ell$ . From this the bounds (8) and (9) can be easily obtained by

using

$$\ell^{-(i-1)} |\hat{x}_i - x_i| \leq |\tilde{\chi}| \leq |\hat{\mathbf{x}} - \mathbf{x}|$$

$$\ell^{-i} |\eta_i - x_{i+1}| \leq |\tilde{\chi}| \leq |\hat{\mathbf{x}} - \mathbf{x}|,$$

which hold for  $\ell \geq 1$ .  $\square$

### A.3 Proof of Proposition 2

Consider the change of coordinates

$$\begin{aligned} \begin{pmatrix} \hat{x}_i \\ \eta_i \end{pmatrix} &\mapsto \tilde{\xi}_i := \begin{pmatrix} \hat{x}_i \\ \eta_i \end{pmatrix} - \begin{pmatrix} x_i \\ x_{i+1} \end{pmatrix} \\ &\quad i = 1, \dots, n-1, \\ \hat{x}_n &\mapsto \tilde{\xi}_n := \hat{x}_n - x_n \end{aligned}$$

that transforms the observer (5)-(6) in the form

$$\dot{\tilde{\xi}} = F\tilde{\xi} + \bar{B}\Delta_\varphi(\tilde{\xi}, x) + G\nu(t) \quad (\text{A.3})$$

where the matrix  $F$  is recursively constructed as

$$F_1 := H_1, \quad F_i := \begin{pmatrix} F_{i-1} & \bar{N}_i \\ \bar{Y}_i & H_i \end{pmatrix}, \quad i = 2, \dots, n-1,$$

and

$$F := \begin{pmatrix} F_{n-1} & 0 \\ \ell\bar{q}_n & -\ell\alpha_n \end{pmatrix}$$

with

$$\begin{aligned} H_i &:= A - D_2(\ell) K_i C \quad \forall i = 1, \dots, n-1, \\ Y_i &:= D_2(\ell) K_i B^T \quad \forall i = 2, \dots, n-1, \end{aligned}$$

the matrix  $G$  is defined as

$$G = \text{col}(G_1, 0, \dots, 0), \quad G_1 := D_2(\ell) K_1,$$

and

$$\Delta_\varphi(\tilde{\xi}, x) := \varphi_s(\Gamma\tilde{\xi} + x) - \varphi(x, d). \quad (\text{A.4})$$

with

$$\Gamma := \text{blkdiag}(\underbrace{C_2, \dots, C_2}_{(n-1) \text{ times}}, 1),$$

By compactly writing the system dynamics (1) as

$$\dot{x} = f(x)$$

the overall dynamics given by the observed system (1), the observer error dynamics (A.3) and the noise generator (4) read as

$$\begin{aligned} \varepsilon\dot{w} &= Sw \\ \dot{x} &= f(x) \\ \dot{\tilde{\xi}} &= F\tilde{\xi} + B\Delta_\varphi(\tilde{\xi}, x) + GPw. \end{aligned} \quad (\text{A.5})$$

Having taken the parameters  $(\underline{\alpha}, \underline{\beta})$  and  $\ell$  according to the prescription of Theorem 1, the trajectories of this system are bounded. The system in question, thus, has a well-defined steady state that can be characterised with the tools proposed in [16]. More specifically, the triangular structure of the system (with the  $x$  and  $w$  subsystem driving the  $\tilde{\xi}$  subsystem) implies the existence of a possibly set-valued function  $\pi_\varepsilon : X \times W \rightrightarrows \mathbb{R}^{2n-1}$  such that the set

$$\text{graph}(\pi_\varepsilon) = \{(w, x, \tilde{\xi}) \in W \times X \times \mathbb{R}^{2n-1} : \tilde{\xi} \in \pi_\varepsilon(w, x)\}$$

is asymptotically stable for (A.5). Furthermore, the properties of the high-gain observer when the measurement noise is absent (*i.e.* when  $w = 0$ ) show that

$$\pi_\varepsilon(0, x) = \{0\} \quad \forall x \in X.$$

The following technical lemma provides an arbitrarily accurate approximation of a continuous selection of  $\pi_\varepsilon(\cdot, \cdot)$ . The lemma refers to a number of functions that enter in definition of the approximation. In order to keep compact the claim of the lemma, we introduce those functions beforehand. In particular, let

$$v := \left\lceil \frac{n}{2} \right\rceil,$$

and let  $\rho$  be an arbitrary (integer) number satisfying  $\rho \geq m$ , with  $m$  given by (10). Note that for any  $n$  we have  $m \geq v$ . The approximation of order  $\rho$  of the steady state is then a function  $\Psi_\varepsilon : W \times X \rightarrow \mathbb{R}^{2n-1}$  defined as

$$\Psi_\varepsilon(w, x) := \text{col}(\Psi_1, \Lambda_1, \Psi_2, \Lambda_2, \dots, \Psi_{n-1}, \Lambda_{n-1}, \Psi_n)$$

in which

$$\begin{aligned} \Psi_i(w, x) &:= \sum_{j=a_i}^{\rho} \psi_{i,j}(w, x) \varepsilon^j, \quad i = 1, \dots, n \\ \Lambda_i(w, x) &:= \sum_{j=b_i}^{\rho} \lambda_{i,j}(w, x) \varepsilon^j, \quad i = 1, \dots, n-1 \end{aligned} \quad (\text{A.6})$$

where the  $a_i$  and  $b_i$  are defined as

$$\begin{aligned} a_i &= i, & i &= 1, \dots, v, \\ a_i &= n - i + 2, & i &= v + 1, \dots, n, \\ b_i &= i, & i &= 1, \dots, v, \\ b_i &= n - i + 1, & i &= v + 1, \dots, n - 1, \end{aligned} \quad (\text{A.7})$$

with

$$\begin{aligned} \psi_{i,j} : X \times W &\rightarrow \mathbb{R}, & i &= 1, \dots, n, \\ & & j &= a_i, \dots, \rho \\ \lambda_{i,j} : X \times W &\rightarrow \mathbb{R}, & i &= 1, \dots, n - 1, \\ & & j &= b_i, \dots, \rho \end{aligned}$$

appropriately defined continuous functions. We have then the following result, instrumental to the proof of Proposition 2.

**Lemma 2** *Consider system (A.5) and the notations introduced before. There exist continuous functions  $\psi_{i,j}(\cdot, \cdot)$  and  $\lambda_{i,j}(\cdot, \cdot)$  such that, having defined*

$$\begin{aligned} E_\varepsilon(w, x) &:= \frac{\partial \Psi_\varepsilon(w, x)}{\partial w} S w + \frac{\partial \Psi_\varepsilon(w, x)}{\partial x} f(x) \\ &\quad - F \Psi_\varepsilon(w, x) - G P w - B \Delta_\varphi(\Psi_\varepsilon(w, x), x), \end{aligned}$$

the following holds

$$\lim_{\varepsilon \rightarrow 0^+} \frac{E_\varepsilon(w, x)}{\varepsilon^{\rho-1}} = 0 \quad \forall (w, x) \in W \times X,$$

$$E_\varepsilon(0, x) = 0 \quad \forall (\varepsilon, x) \in [0, 1] \times X.$$

Furthermore, there exist continuous functions  $\bar{\psi}_{i,a_i}(\cdot, \cdot)$ ,  $i = 1, \dots, n$ , satisfying

$$\begin{aligned} \psi_{i,a_i}(w, x) &:= \ell^{2i-1} \bar{\psi}_{i,a_i}(w, x), & i &= 1, \dots, m, \\ \psi_{i,a_i}(w, x) &:= \ell \bar{\psi}_{i,a_i}(w, x), & i &= m + 1, \dots, n. \end{aligned}$$

**Proof:** First of all consider the case where  $w = 0$ . By recalling the definition of  $\Delta_\varphi(\cdot, \cdot)$  in (A.4) it is easy to verify that  $\Psi_\varepsilon(0, x) = 0$  makes  $E_\varepsilon(0, x) = 0$ . As a consequence in the following we will show that  $\Psi_\varepsilon(w, x)$  can be chosen as a continuous function in  $w$  satisfying  $\Psi_\varepsilon(0, x) = 0$ .

Now, since  $w$  and  $x$  range in bounded sets and the function  $\psi_{i,j}(\cdot, \cdot)$  and  $\lambda_{i,j}(\cdot, \cdot)$  are continuous, we have that

$$\lim_{\varepsilon \rightarrow 0^+} \Psi_\varepsilon(w, x) = 0 \quad \forall (w, x) \in W \times X.$$

Let us denote  $\bar{\Psi}_\varepsilon = \Gamma \Psi_\varepsilon$  and let denote with  $\bar{\Psi}_i$  the  $i$ -th element of  $\bar{\Psi}_\varepsilon$ . Expanding  $\Delta_\varphi(\bar{\Psi}_\varepsilon, x)$  by a Taylor series around  $\bar{\Psi}_\varepsilon = 0$  we obtain

$$\Delta_\varphi(\bar{\Psi}_\varepsilon, x) = \sum_{i=1}^{\rho} \varphi_i(x) [\bar{\Psi}_\varepsilon]^i + \tau_\rho(\bar{\Psi}_\varepsilon, x)$$

in which  $\varphi_i(\cdot)$ ,  $i = 1, \dots, \rho$ , are properly defined continuous functions,  $\tau_\rho(\cdot, \cdot)$  is a properly defined continuous remainder function, and the  $[\bar{\Psi}_\varepsilon]^i$  are monomials of the form

$$[\bar{\Psi}_\varepsilon]^i = \prod_{j=1}^{n-1} \bar{\Psi}_j^{k_j}, \quad \sum_{j=1}^{n-1} k_j = i.$$

By replacing the  $\Psi_i$  in the definition of  $\bar{\Psi}_i$  with the expression (A.6) and grouping the terms with the same power of  $\varepsilon$ , the Taylor series expansion of  $\Delta_\varphi(\cdot, \cdot)$  can be rewritten as

$$\Delta_\varphi(\bar{\Psi}_\varepsilon, x) = \sum_{j=1}^{\rho} \varepsilon^j \phi_j(w, x) + \varepsilon^{\rho+1} R_\varepsilon(w, x) \quad (\text{A.9})$$

where the functions  $\phi_j(\cdot, \cdot)$ ,  $j = 1, \dots, \rho$ , and  $R_\varepsilon(\cdot, \cdot)$  are appropriately defined continuous functions satisfying  $\phi_j(0, x) = 0$  and  $R_\varepsilon(0, x) = 0$ . As far as the  $\phi_j$ 's are concerned, in particular, we note that, because the  $\bar{\Psi}_i$  are polynomials in  $\varepsilon$  and the  $[\bar{\Psi}_\varepsilon]^j$  are polynomials in the  $\bar{\Psi}_i$ , only the coefficients of power smaller or equal to  $j$  in  $\varepsilon$  in the  $\bar{\Psi}_i$  can be in  $\phi_j$ . Namely  $\phi_j(\cdot, \cdot)$  depends only on  $\psi_{i,k}$  with  $k \leq i$  for all  $i = 1, \dots, n$  and  $j = 1, \dots, \rho$ .

Consider now the expression of  $E_\varepsilon(\cdot, \cdot)$  and, by letting

$$E_\varepsilon(\cdot, \cdot) := \text{col} \left( E_1, \Xi_1, E_2, \Xi_2, \dots, E_{n-1}, \Xi_{n-1}, E_n \right)$$

note that

$$\begin{aligned} E_1 &= \dot{\Psi}_1 + \ell \alpha_1 \Psi_1 - \Lambda_1 - \ell \alpha_1 P w \\ \Xi_1 &= \dot{\Lambda}_1 + \ell^2 \beta_1 \Psi_1 - \Lambda_2 - \ell^2 \beta_1 P w \\ &\vdots \\ E_i &= \dot{\Psi}_i + \ell \alpha_i \Psi_i - \Lambda_i - \ell \alpha_i \Lambda_{i-1} \\ \Xi_i &= \dot{\Lambda}_i + \ell^2 \beta_i \Psi_i - \Lambda_{i+1} - \ell^2 \beta_i \Lambda_{i-1} \\ &\vdots \\ E_{n-1} &= \dot{\Psi}_{n-1} + \ell \alpha_{n-1} \Psi_{n-1} - \Lambda_{n-1} - \ell \alpha_{n-1} \Lambda_{n-2} \\ \Xi_{n-1} &= \dot{\Lambda}_{n-1} + \ell^2 \beta_{n-1} \Psi_{n-1} - \Delta_\varphi(\bar{\Psi}_\varepsilon, x) \\ &\quad - \ell^2 \beta_{n-1} \Lambda_{n-2} \\ E_n &= \dot{\Psi}_n + \ell \alpha_n \Psi_n - \Delta_\varphi(\bar{\Psi}_\varepsilon, x) - \ell \alpha_n \Lambda_{n-1} \end{aligned}$$

$$\begin{aligned}
E_1 &= \sum_{j=1}^{\rho} \left[ L_f \psi_{1,j} + \frac{1}{\varepsilon} L_S \psi_{1,j} \right] \varepsilon^j + \ell \alpha_1 \sum_{j=1}^{\rho} \psi_{1,j} \varepsilon^j - \sum_{j=1}^{\rho} \lambda_{1,j} \varepsilon^j - \ell \alpha_1 P w \\
E_i &= \sum_{j=a_i}^{\rho} \left[ L_f \psi_{i,j} + \frac{1}{\varepsilon} L_S \psi_{i,j} \right] \varepsilon^j + \ell \alpha_i \sum_{j=a_i}^{\rho} \psi_{i,j} \varepsilon^j - \sum_{j=b_i}^{\rho} \lambda_{i,j} \varepsilon^j - \ell \alpha_i \sum_{j=b_{i-1}}^{\rho} \lambda_{i-1,j} \varepsilon^j \\
E_n &= \sum_{j=2}^{\rho} \left[ L_f \psi_{n,j} + \frac{1}{\varepsilon} L_S \psi_{n,j} \right] \varepsilon^j + \ell \alpha_n \sum_{j=a_n}^{\rho} \psi_{n,j} \varepsilon^j - \sum_{j=1}^{\rho} \varepsilon^j \phi_j - \varepsilon^{\rho+1} R_\varepsilon - \ell \alpha_n \sum_{j=b_{n-1}}^{\rho} \lambda_{n-1,j} \varepsilon^j \\
\Xi_1 &= \sum_{j=1}^{\rho} \left[ L_f \lambda_{1,j} + \frac{1}{\varepsilon} L_S \lambda_{1,j} \right] \varepsilon^j + \ell^2 \beta_2 \sum_{j=1}^{\rho} \psi_{1,j} \varepsilon^j - \sum_{j=2}^{\rho} \lambda_{2,j} \varepsilon^j - \ell^2 \beta_2 P w \\
\Xi_i &= \sum_{j=b_i}^{\rho} \left[ L_f \lambda_{i,j} + \frac{1}{\varepsilon} L_S \lambda_{i,j} \right] \varepsilon^j + \ell^2 \beta_i \sum_{j=a_i}^{\rho} \psi_{i,j} \varepsilon^j - \sum_{j=b_{i+1}}^{\rho} \lambda_{i+1,j} \varepsilon^j - \ell^2 \beta_i \sum_{j=b_{i-1}}^{\rho} \lambda_{i-1,j} \varepsilon^j \\
\Xi_{n-1} &= \sum_{j=2}^{\rho} \left[ L_f \lambda_{n-1,j} + \frac{1}{\varepsilon} L_S \lambda_{n-1,j} \right] \varepsilon^j + \ell^2 \beta_{n-1} \sum_{j=a_{n-1}}^{\rho} \psi_{n-1,j} \varepsilon^j - \sum_{j=1}^{\rho} \varepsilon^j \phi_j - \varepsilon^{\rho+1} R_\varepsilon - \ell^2 \beta_{n-1} \sum_{j=b_{n-2}}^{\rho} \lambda_{n-2,j} \varepsilon^j
\end{aligned} \tag{A.8}$$

where, for the sake of compactness, we omitted the argument  $(w, x)$  from the functions  $\Psi_i$ ,  $i = 1, \dots, n$ ,  $\Lambda_i$ ,  $i = 1, \dots, n-1$  and  $\overline{\Psi}_\varepsilon$ .

By embedding (A.6) and (A.9) in the previous expressions, we obtain the expressions in (A.8) in which

$$\begin{aligned}
L_f \psi_{i,j} &:= \frac{\partial \psi_{i,j}(w, x)}{\partial x} f(x), \quad L_S \psi_{i,j} := \frac{\partial \psi_{i,j}(w, x)}{\partial w} S w, \\
L_f \lambda_{i,j} &:= \frac{\partial \lambda_{i,j}(w, x)}{\partial x} f(x), \quad L_S \lambda_{i,j} := \frac{\partial \lambda_{i,j}(w, x)}{\partial w} S w.
\end{aligned}$$

The idea now is to iteratively select the functions  $\psi_{i,j+1}(\cdot, \cdot)$ ,  $\lambda_{i,j+1}(\cdot, \cdot)$  to annihilate, in the previous expressions, the terms in  $\varepsilon$  of order  $j$ , with  $j = 0, \dots, \rho-1$ , for  $i = 1, \dots, n$ . We start by considering the term of order 0 in  $\varepsilon$  in the expression of  $E_1$  and  $\Xi_1$  which are annihilated by taking

$$\begin{aligned}
\psi_{1,1}(w, x) &= \ell \alpha_1 P S^{-1} w, \\
\lambda_{1,1}(w, x) &= \ell^2 \beta_1 P S^{-1} w.
\end{aligned} \tag{A.10}$$

We observe that  $\psi_{1,1}(w, x)$ ,  $\lambda_{1,1}(w, x)$ , are polynomials in  $w$  of order 1 with constant coefficients. Since  $a_i = b_i = i$  for  $i = 1, \dots, v$ , having fixed the terms  $\psi_{1,1}$  and  $\lambda_{1,1}$  it is possible, iteratively, to select all the functions  $\psi_{i,i}(w, x)$  and  $\lambda_{i,i}(w, x)$  for  $i = 2, \dots, v-1$  to annihilate the terms in  $\varepsilon$  of order  $i-1$  in  $E_i$  and  $\Xi_i$  by solving the following PDEs

$$\begin{aligned}
-L_S \psi_{i,i} &= \ell \alpha_i \lambda_{i-1,i-1} \\
-L_S \lambda_{i,i} &= \ell^2 \beta_i \lambda_{i-1,i-1}.
\end{aligned}$$

Using the fact that  $S$  is invertible the previous PDEs admit a solution which is polynomial in  $w$  of order 1 with

constant coefficients. With this we have fixed the terms  $\psi_{i,i}$  and  $\lambda_{i,i}$  for  $i = 1, \dots, v-1$ . Once the terms above have been fixed, it is possible to repeat the selection process for  $j = 1, \dots, v-1$  from the top in order to select the functions  $\psi_{i,j+i}$  and  $\lambda_{i,j+i}$  for  $i = 1, \dots, v-j-1$ , to annihilate the terms in  $\varepsilon$  of order  $j+i-1$  in  $E_i$  and  $\Xi_i$ , by solving the following PDEs

$$\begin{aligned}
-L_S \psi_{1,1+j} &= \ell \alpha_1 \psi_{1,j} - \lambda_{1,j} \\
-L_S \lambda_{1,1+j} &= \ell^2 \beta_2 \psi_{1,j} - \lambda_{2,j}
\end{aligned}$$

and

$$\begin{aligned}
-L_S \psi_{i,j+i} &= \ell \alpha_i (\psi_{i,j+i-1} - \lambda_{i-1,j+i-1}) - \lambda_{i,j+i-1} \\
-L_S \lambda_{i,j+i} &= \ell^2 \beta_i (\psi_{i,j+i-1} - \lambda_{i-1,j+i-1}) - \lambda_{i+1,j+i-1}
\end{aligned}$$

for  $i = 2, \dots, v-j-1$ . Again, using the fact that  $S$  is invertible the previous PDEs admit a solution which is polynomial in  $w$  of order 1 with constant coefficients. With this we have assigned all the terms  $\psi_{i,j}$  and  $\lambda_{i,j}$  for  $i = 1, \dots, v-1$  and  $j = i, \dots, v$ .

Now let us consider the terms  $\Xi_i$  for  $i = v+1, \dots, n-1$  by starting from the bottom. Note that  $\phi_1(\cdot, \cdot)$  depends only on  $\psi_{1,1}$ . Hence we can assume that  $\phi_1(w, x)$  is a polynomial in  $w$  of order 1 with coefficients dependent on  $x$  and vanishing when  $w = 0$ . As a consequence the terms in  $\varepsilon$  of order 1 in  $\Xi_{n-1}$  are annihilated if  $\lambda_{n-1,2}$  can be chosen such that

$$L_S \lambda_{n-1,2} = \phi_1. \tag{A.11}$$

Using the fact that  $S$  is invertible and function  $\phi_1$  is a polynomial in  $w$  the previous PDE admits a solution

which is polynomial of order 1 in  $w$  with coefficients which depends on  $x$ . Recall that for  $i = v+1, \dots, n-2$  we have  $b_i > b_{i+1}$ . As a consequence, once the term  $\lambda_{n-1,2}$  has been fixed we can proceed iteratively in order to select (from the bottom) all the functions  $\lambda_{i,b_i}(w, x)$  for  $i = v+1, \dots, n-2$  by annihilating the terms in  $\varepsilon$  of order  $b_i - 1$  in  $\Xi_i$  by solving the following PDEs

$$L_S \lambda_{i,b_i} = \lambda_{i+1,b_{i+1}}.$$

Using the fact that  $S$  is invertible the previous PDEs admit a solution which is polynomial of order 1 in  $w$  with coefficients which depends on  $x$ . With this we have fixed all the terms  $\lambda_{i,b_i}$  for  $i = v+1, \dots, n-1$ .

Now consider the terms  $E_i$  for  $i = v+1, \dots, n$  by starting from the bottom. The terms in  $\varepsilon$  of order 1 in  $E_n$  are annihilated if  $\psi_{n,2}$  can be chosen such that

$$L_S \psi_{n,2} = \phi_1. \quad (\text{A.12})$$

As said for (A.11), the previous PDE admits a solution which is polynomial of order 1 in  $w$  with coefficients that depends on  $x$ . Concerning the terms  $E_i$  for  $i = v+1, \dots, n-1$ , once the terms  $\lambda_{i,b_i}$  has been fixed for  $i = v+1, \dots, n-1$ , we can proceed iteratively in order to select the functions  $\psi_{i,a_i}(w, x)$  by annihilating the terms in  $\varepsilon$  of order  $a_i - 1$  in  $E_i$  by solving the following PDEs

$$L_S \psi_{i,a_i} = \lambda_{i,b_i}.$$

Using the fact that  $S$  is invertible the previous PDEs admit a solution which is polynomial of order 1 in  $w$  with coefficients which depends on  $x$ . With this we have fixed all the terms  $\psi_{i,a_i}$  for  $i = v+1, \dots, n$ .

We note that the function  $\phi_2$  depends only on the terms  $\psi_{i,j}$ ,  $i = 1, \dots, n$ , with powers of  $\varepsilon$  smaller (or equal) to two, namely  $j \leq 2$ , and therefore on the functions  $\psi_{1,1}$ ,  $\psi_{1,2}$ ,  $\psi_{2,2}$  and  $\psi_{n,2}$  which have already been fixed. As a consequence, the previous arguments can be used in order to select also the function  $\lambda_{n-1,3}$  by annihilating the terms in  $\varepsilon$  of order 2 in  $\Xi_{n-1}$  as solution of

$$L_S \lambda_{n-1,3} = \phi_2,$$

and then, by induction, the functions  $\lambda_{i,b_{i+1}}$  by annihilating the terms in  $\varepsilon$  of order  $b_i$  in  $\Xi_i$  for  $i = v+1, \dots, n-2$  (from the bottom), as solution of

$$L_S \lambda_{i,b_{i+1}} = \lambda_{i+1,b_{i+1}}.$$

Now note that the term  $\phi_k$  depends on the terms  $\psi_{i,j}$ , with  $i \in [1, n]$  and  $a_i \leq j \leq k$ . As a consequence, once the terms  $\lambda_{i,b_{i+1}}$  have been fixed for  $i = v+1, \dots, n-1$ , this procedure can be generalized at the following steps, in order to select all the functions  $\lambda_{i,j}$  for  $i =$

$v+1, \dots, n-1$  and  $j = b_i+1, \dots, v-1$  by solving, from the bottom, the following PDEs

$$\begin{aligned} L_S \lambda_{i,j+1} &= \lambda_{i+1,j} + \ell^2 \beta_i (\lambda_{i-1,j} - \psi_{i,j}) \\ L_S \lambda_{n-1,j+1} &= \phi_j + \ell^2 \beta_{n-1} (\lambda_{n-2,j} - \psi_{n-1,j}) \end{aligned}$$

and the functions  $\psi_{i,j}$ , for  $i = v+1, \dots, n$  and  $j = a_i, \dots, v-1$  by solving, from the bottom, the following PDEs

$$\begin{aligned} L_S \psi_{i,j+1} &= \lambda_{i,j} + \ell \alpha_i (\lambda_{i-1,j} - \psi_{i,j}) \\ L_S \psi_{n,j+1} &= \phi_j + \ell \alpha_n (\lambda_{n-1,j} - \psi_{n,j}) \end{aligned}$$

where we considered

$$\psi_{i,j} = 0 \quad \forall j < a_i, \quad \lambda_{i,j} = 0 \quad \forall j < b_i.$$

With this we have fixed all the terms  $\psi_{i,j}$  for  $i = v+1, \dots, n$ ,  $j = a_i, \dots, v$  and  $\lambda_{i,j}$ , for  $i = v+1, \dots, n-1$ ,  $j = b_i, \dots, v$ ,

As far as the terms  $\Xi_v$  and  $E_v$  are concerned, we can select the functions  $\psi_{v,a_v}$  and  $\lambda_{v,b_v}$  by solving the following PDEs

$$\begin{aligned} L_S \psi_{v,v} &= \ell \alpha_v \lambda_{v-1,v-1} \\ L_S \lambda_{v,v} &= \ell^2 \beta_v \lambda_{v-1,v-1} + \lambda_{v+1,b_{v+1}}. \end{aligned}$$

Since the term  $\lambda_{v+1,b_{v+1}}$  has been already fixed by the previous procedure, by using the fact that  $S$  is invertible and functions  $\lambda_{v-1,v-1}$  and  $\lambda_{v+1,b_{v+1}}$  are polynomial in  $w$ , the previous PDEs admit a solution which is polynomial of in  $w$  with coefficients which may depends on  $x$ . With this we have finally assigned all the terms  $\psi_{i,j}$ , for  $i = 1, \dots, n$ ,  $j = a_i, \dots, v$  and  $\lambda_{i,j}$  for  $i = 1, \dots, n-1$  and  $j = a_i, \dots, v$ . Evidently, the iterative procedure may be lead similarly in order to select all the remaining terms  $\psi_{i,j+1}$ , for  $i = 1, \dots, n$ ,  $j = v, \dots, \rho$  and  $\lambda_{i,j+1}$  for  $i = 1, \dots, n-1$  and  $j = v, \dots, \rho$  by annihilating at each step the terms in  $\varepsilon$  of order  $j$  which depends on the terms  $\psi_{i,j}$ ,  $\lambda_{i,j}$  which have already been fixed in the previous iterative process.

Finally, by embedding those functions in the expressions of  $E_i(\cdot, \cdot)$ ,  $i = 1, \dots, n$  and  $\Xi_i(\cdot, \cdot)$ ,  $i = 1, \dots, n-1$  and bearing in mind the definition of  $R_\varepsilon(\cdot, \cdot)$ , it is readily

seen that

$$\begin{aligned} E_1(w, x) &= \varepsilon^\rho [L_f \psi_{1,\rho} + \ell \alpha_1 \psi_{1,\rho} - \lambda_{1,\rho}] \\ &\vdots \\ E_i(w, x) &= \varepsilon^\rho [L_f \psi_{i,\rho} + \ell \alpha_i (\psi_{i,\rho} - \lambda_{i-1,\rho}) - \lambda_{i,\rho}] \\ &\vdots \\ E_n(w, x) &= \varepsilon^\rho [L_f \psi_{n,\rho} + \ell \alpha_n (\psi_{n,\rho} - \lambda_{n-1,\rho}) - \phi_\rho] \\ &\quad + \varepsilon^{\rho+1} w \bar{R}_\varepsilon(w, x) \end{aligned}$$

and

$$\begin{aligned} \Xi_1(w, x) &= \varepsilon^\rho [L_f \lambda_{1,\rho} + \ell^2 \beta_1 \psi_{1,\rho} - \lambda_{2,\rho}] \\ &\vdots \\ \Xi_i(w, x) &= \varepsilon^\rho [L_f \lambda_{i,\rho} + \ell^2 \beta_i (\psi_{i,\rho} - \lambda_{i-1,\rho}) - \lambda_{i+1,\rho}] \\ &\vdots \\ \Xi_{n-1}(w, x) &= \varepsilon^\rho [L_f \lambda_{n-1,\rho} + \ell^2 \beta_{n-1} (\psi_{n-1,\rho} - \lambda_{n-2,\rho}) \\ &\quad - \phi_\rho] + \varepsilon^{\rho+1} |w| \bar{R}_\varepsilon(w, x) \end{aligned}$$

where  $\bar{R}_\varepsilon(\cdot, \cdot)$  is an appropriately defined continuous function. By using the previous expressions and the fact that  $\psi_{i,j}(0, x) = 0$ ,  $\lambda_{i,j}(0, x) = 0$  for any  $i = 1, \dots, n-1$  and  $j = 1, \dots, \rho$ , the first part of the lemma immediately follows.

Now, by using (A.10), we have for  $i = 1$

$$\psi_{1,1}(w, x) := \ell \alpha_1 P S^{-1} w = \ell \bar{\psi}_{1,1}(w, x),$$

For  $i = 2, \dots, m-1$  we have  $a_{i-1} < a_j$ . As a consequence the terms  $\psi_{i,a_i}(w, x)$  are computed a solution of of the PDEs

$$L_S \psi_{i,a_i} := \ell \alpha_i \lambda_{i-1,b_{i-1}}$$

namely they depend on the terms  $\lambda_{i-1,b_{i-1}}$ . Therefore, by using (A.10), we have that  $\lambda_{1,1}$  is chosen as

$$\lambda_{1,1}(w, x) = \ell^2 \beta_1 P S^{-1} w := \ell^2 \bar{\lambda}_{1,1}(w, x).$$

For  $i = 2, \dots, m-1$ , since  $b_{j-1} < b_{j+1}$ , the  $\lambda_{i,b_i}$  are computed as solution of

$$L_S \lambda_{i,b_i} = \ell^2 \beta_i \lambda_{i,b_{i-1}}.$$

With this in mind, we can always select functions  $\bar{\psi}_{i,i}(w, x)$  and  $\bar{\lambda}_{i,i}(w, x)$  such that

$$\begin{aligned} \psi_{i,a_i}(w, x) &:= \ell^{2i-1} \bar{\psi}_{i,a_i}(w, x) \quad i = 1, \dots, m-1. \\ \lambda_{i,b_i}(w, x) &:= \ell^{2i} \bar{\lambda}_{i,b_i}(w, x) \end{aligned}$$

Now we start from the bottom by computing  $\psi_{n,2}$ . It is chosen as solution of (A.12). As a consequence, since  $\phi_1$  depends only on  $\psi_{1,1}$  which is a function in  $\ell$ , we can deduce that  $\psi_{n,2}$  can be written as

$$\psi_{n,2}(w, x) := \ell \bar{\psi}_{n,2}(w, x).$$

Also, for  $i = m+1, \dots, n-1$  we have that  $a_i < a_{i-1}$ . As a consequence the functions  $\psi_{i,a_i}$  are chosen as solution of

$$L_S \psi_{i,a_i} = \lambda_{i,b_i}.$$

Therefore we need to compute before the terms  $\lambda_{i,b_i}$ . The term  $\lambda_{n-1,2}$  is computed as solution of (A.11). By repeating the previous arguments, we can deduce that also  $\lambda_{n,2}$  is a term in  $\ell$ , namely it can be written as

$$\lambda_{n-1,2}(w, x) := \ell \bar{\lambda}_{n-1,2}(w, x).$$

For  $i = m+1, \dots, n-2$  we can compute  $\lambda_{i,b_i}$  as solution of

$$L_S \lambda_{i,b_i} = \lambda_{i+1,b_{i+1}}$$

because  $b_{i+1} < b_{i-1}$ . As a consequence the  $\lambda_{i,b_i}$  are all functions of  $\lambda_{n-1,2}$  which is a term in  $\ell$ . By induction, we conclude that

$$\begin{aligned} \psi_{i,a_i}(w, x) &= \ell \bar{\psi}_{i,a_i}(w, x) \\ \lambda_{i,b_i}(w, x) &= \ell \bar{\lambda}_{i,b_i}(w, x) \end{aligned} \quad i = m+1, \dots, n-1$$

for some  $\bar{\psi}_{i,a_i}(w, x)$  and  $\bar{\lambda}_{i,b_i}(w, x)$ . It remains to fix the terms  $\psi_{b,a_m}$ ,  $\lambda_{m,b_m}$ .

If  $n$  is odd we have  $b_{m-1} < b_m$ . As a consequence  $\psi_m$  can be computed as solution of

$$L_S \psi_{m,a_m} = \ell \alpha_m \lambda_{m-1,b_{m-1}}.$$

But being  $\lambda_{m-1,b_{m-1}}$  a term in  $\ell^{2(m-1)}$  we get that

$$\psi_{m,a_m}(x, w) := \ell^{2m-1} \bar{\psi}_{m,a_m}(x, w)$$

for some  $\bar{\psi}_{m,a_m}(x, w)$ .

If  $n$  is even we have  $b_{m+1} < b_m = b_{m-1}$ . As a consequence  $\lambda_{m,b_m}$  is chosen as solution of

$$L_S \lambda_{m,b_m} = \lambda_{m+1,b_{m+1}}$$

resulting in a term in  $\ell$ , whereas  $\psi_{m,a_m}$  is chosen as solution of

$$L_S \psi_{m,a_m} = \ell \alpha_m \lambda_{m-1,b_{m-1}} + \lambda_{m,b_m}.$$

But since  $\lambda_{m,b_m}$  is a term in  $\ell$  whereas  $\lambda_{m-1,b_{m-1}}$  is a term in  $\ell^{2(m-1)}$  we get again

$$\psi_{m,a_m}(x,w) := \ell^{2m-1} \bar{\psi}_{m,a_m}(x,w)$$

for some  $\bar{\psi}_{m,a_m}(x,w)$ . This concludes the proof of Lemma 2.  $\square$

With the result of Lemma 2 in hand, we are now in the position of proving Proposition 2. Let consider the change of variables

$$\tilde{\xi} \mapsto \eta := \tilde{\xi} - \Psi_\varepsilon(w, x),$$

with  $\Psi_\varepsilon(\cdot, \cdot)$  introduced in the previous lemma with a  $\rho > 1$  and note that, by bearing in mind the definition of  $E_\varepsilon(\cdot, \cdot)$ ,

$$\dot{\Psi}_\varepsilon = F\Psi_\varepsilon + B\Delta_\varphi(\Psi_\varepsilon, x) + GPw + E_\varepsilon(w, x).$$

Furthermore, note that

$$\begin{aligned} \Delta_\varphi(\tilde{\xi}, x) - \Delta_\varphi(\Psi_\varepsilon(w, x), x) &= \Delta_\varphi(\eta + \Psi_\varepsilon(w, x), x) - \Delta_\varphi(\Psi_\varepsilon(w, x), x) \\ &= \varphi_s(\eta + \Psi_\varepsilon(w, x) + x) - \varphi(x) \\ &\quad - (\varphi_s(\Psi_\varepsilon(w, x) + x) - \varphi(x)) \\ &= \varphi_s(\eta + \Psi_\varepsilon(w, x) + x) - \varphi_s(\Psi_\varepsilon(w, x) + x) \\ &= \Delta_\varphi(\eta, \Psi_\varepsilon + x), \end{aligned}$$

and that there exists a  $\varepsilon_1^*(\ell) \in (0, 1]$  such that for all positive (the value of  $\varepsilon^*$  depends on, among other things, the choice of the set  $X_\delta$  on which  $\varphi_s(\cdot)$  coincides with  $\varphi(\cdot)$ )  $\varepsilon \leq \varepsilon_1^*(\ell)$

$$\Delta_\varphi(0, \Psi_\varepsilon(w, x) + x) = 0 \quad \forall (w, x) \in W \times X.$$

By the previous facts the error dynamics in the new coordinates can be easily computed as

$$\dot{\eta} = F\eta + B\Delta_\varphi(\eta, \Psi_\varepsilon(w, x) + x) + E_\varepsilon(w, x). \quad (\text{A.13})$$

Since the Lipschitz constant of  $\Delta_\varphi(\cdot, \cdot)$  is not affected by the value of the arguments, the same values of  $\ell$  that make system (A.3) ISS with respect to the input  $\nu(t)$  make also system (A.13) ISS with respect to the input  $E_\varepsilon(\cdot, \cdot)$ . In particular, there exists a positive constant  $c_0$  such that

$$\begin{aligned} \limsup_{t \rightarrow \infty} |\eta(t)| &= \limsup_{t \rightarrow \infty} |\tilde{\xi}(t) - \Psi_\varepsilon(w(t), x(t))| \\ &\leq c_0 \limsup_{t \rightarrow \infty} |E_\varepsilon(w(t), x(t))| \\ &\leq c_0 \|E_\varepsilon(w(\cdot), x(\cdot))\|_\infty \end{aligned}$$

Using the fact that, for any  $\rho \geq m$ ,  $E_\varepsilon(w, x)$  is a term in  $\varepsilon^\rho$ , it follows that there exists a positive constant  $c_1$  such that

$$\limsup_{t \rightarrow \infty} |\eta(t)| \leq c_1 \varepsilon^\rho \|w(\cdot)\|_\infty.$$

Consider now the the expressions of the components  $\Psi_i(\cdot, \cdot)$ ,  $i = 1, \dots, n$ , of  $\Psi_\varepsilon(\cdot, \cdot)$  introduced in Lemma 2. It turns out that there exist a positive  $\varepsilon_2^*(\ell) \leq \varepsilon_1^*(\ell)$  and a positive constant  $\mu_2$  such that

$$|\Psi_i(w, x)| \leq c_2 \varepsilon^i \ell^{2i-1} |w| \quad i = 1, \dots, m,$$

$$|\Psi_i(w, x)| \leq c_2 \varepsilon^{n-j+2} \ell |w| \quad i = m+1, \dots, n,$$

for all positive  $\varepsilon \leq \varepsilon_2^*(\ell)$  and for all  $(w, x) \in W \times X$ . From this, for all  $j = 1, \dots, m$ , we have

$$\begin{aligned} \limsup_{t \rightarrow \infty} |\hat{x}_i(t) - x_i(t)| &= \limsup_{t \rightarrow \infty} |\eta_{i1}(t) + \Psi_i(w(t), x(t))| \\ &\leq \limsup_{t \rightarrow \infty} |\eta_{i+1}(t)| + \|\Psi_i(w(\cdot), x(\cdot))\|_\infty \\ &\leq c_1 \varepsilon^\rho \|w(\cdot)\|_\infty + c_2 \varepsilon^i \ell^{2i-1} \|w(\cdot)\|_\infty \end{aligned}$$

and for  $i = m+1, \dots, n$  we have

$$\begin{aligned} \limsup_{t \rightarrow \infty} |\hat{x}_i(t) - x_i(t)| &= \lim_{t \rightarrow \infty} \sup |\eta_{i1}(t) + \Psi_i(w(t), x(t))| \\ &\leq \limsup_{t \rightarrow \infty} |\eta_i(t)| + \|\Psi_i(w(\cdot), x(\cdot))\|_\infty \\ &\leq c_1 \varepsilon^\rho \|w(\cdot)\|_\infty + c_2 \varepsilon^{n-i+2} \ell \|w(\cdot)\|_\infty. \end{aligned}$$

Since  $\nu(t) = Pw(t)$ , and by recalling that  $\|w(\cdot)\|_\infty$  does not depend on the choice of  $\varepsilon$ , there exists a  $c_3$  such satisfying

$$\|w(\cdot)\|_\infty \leq c_3 \|\nu(\cdot)\|_\infty.$$

The result follows by taking an appropriate  $\varepsilon^*(\ell) \leq \varepsilon_2^*(\ell)$  and  $c > 0$ .  $\square$

#### A.4 Proof of Proposition 3

Consider the change of coordinates

$$\begin{aligned} \begin{pmatrix} \hat{x}_i \\ \eta_i \end{pmatrix} &\mapsto \zeta_i := \begin{pmatrix} \zeta_{i1} \\ \zeta_{i2} \end{pmatrix} := \begin{pmatrix} \hat{x}_i - x_i \\ \ell^{-1}(\eta_i - x_{i+1}) \end{pmatrix} \\ &\quad i = 1, \dots, n-1, \\ \hat{x}_n &\mapsto \zeta_n := (\hat{x}_n - x_n), \end{aligned} \quad (\text{A.14})$$

that transforms system (11)-(12) into

$$\begin{aligned}\dot{\zeta}_1 &= \ell E_1 \zeta_1 + \ell^{-1} B_2 u_1 + \ell K_1 \nu \\ \dot{\zeta}_i &= \ell E_i \zeta_i + \ell^{-1} B_2 u_i + \ell K_i \text{sat}_{r_i}(\ell B_2^T \zeta_{i-1}) \\ &\quad i = 2, \dots, n-1 \\ \dot{\zeta}_n &= -\ell \alpha_n \zeta_n + u_{n-1} + \ell \alpha_n \text{sat}_{r_n}(\ell B_2^T \zeta_{n-1})\end{aligned}$$

where the variables  $u_i$ ,  $i = 1, \dots, n-1$ , are defined as

$$\begin{aligned}u_i &:= \text{sat}_{r_{i+2}}(\eta_{i+1}) - x_{i+2}, \quad i = 1, \dots, n-2 \\ u_{n-1} &:= \varphi_s(\hat{x}, 0) - \varphi(x, d).\end{aligned}$$

By definition of saturation function, of  $\varphi_s(\cdot, \cdot)$  and since  $x(t)$  and  $d(t)$  range in compact sets, there exist  $\bar{u}_i > 0$ ,  $i = 1, \dots, n-1$ , independent of  $\ell$ , such that

$$\|u_i(\cdot)\|_\infty \leq \bar{u}_i \quad \forall i = 1, \dots, n-1. \quad (\text{A.15})$$

Furthermore, note that, by definition of  $\mathcal{S}'$ , the matrices  $E_i$ ,  $i = 1, \dots, n-1$ , are Hurwitz and  $\alpha_n > 0$ . Hence, by applying Lemma 3 in Appendix A.5, it turns out that there exist positive constants  $c_{ij}$ , with  $i = 1, \dots, n$  and  $j = 1, \dots, 4$ , such that, for any  $\ell \geq 1$ , we get

$$\begin{aligned}|\zeta_1(t)| &\leq c_{11} \exp(-c_{12} \ell t) |\zeta_1(0)| + \frac{c_{13}}{\ell^2} \bar{u}_1 + c_{14} \|\nu(\cdot)\|_\infty \\ |\zeta_i(t)| &\leq c_{i1} \exp(-c_{i2} \ell t) |\zeta_i(0)| + \frac{c_{i3}}{\ell^2} \bar{u}_i + c_{i4} \\ &\quad i = 2, \dots, n.\end{aligned}$$

From this, by using the fact that for all  $\ell \geq 1$

$$\begin{aligned}|x_i - \hat{x}_i| &\leq |\zeta_i| \leq |x_i - \hat{x}_i| + |x_{i+1} - \eta_i|, \\ \frac{1}{\ell} |x_{i+1} - \eta_i| &\leq |\zeta_i| \leq |x_i - \hat{x}_i| + |x_{i+1} - \eta_i|\end{aligned}$$

the result in item (a) of the proposition immediately follows with  $\bar{p}_i$ ,  $i = 1, \dots, n$ , defined as

$$\begin{aligned}\bar{p}_1 &:= c_{11} \pi_1 + c_{13} \bar{u}_1 + c_{14} \bar{\nu}, \\ \bar{p}_i &:= c_{i1} \pi_i + c_{i3} \bar{u}_i + c_{i4}, \quad i = 2, \dots, n\end{aligned}$$

where

$$\pi_i := \max_{x \in X, (\hat{x}, \eta) \in \widehat{X} \times E} \{|x_i - \hat{x}_i| + |x_{i+1} - \eta_i|\}, \quad i = 1, \dots, n-1$$

$$\pi_n := \max_{x \in X, \hat{x} \in \widehat{X}} |x_n - \hat{x}_n|.$$

To prove item (b) we proceed by induction by recursively showing that all the  $\eta_i$ , from  $i = 1$  to  $i = n-1$ , exit from the saturation if the measurement noise is sufficiently small.

**Case  $i = 1$ .** Consider the change of coordinates (A.2) and the first component  $\tilde{\chi}_1$  whose dynamics are given by

$$\dot{\tilde{\chi}}_1 = \ell E_1 \tilde{\chi}_1 + \ell^{-1} B_2 u_1 + \ell K_1 \nu(t).$$

We observe that the initial condition  $\tilde{\chi}_1(0)$  ranges in a compact set  $\mathcal{E}_1$  not dependent on  $\ell$  (for all  $\ell \geq 1$ ). Using the fact that  $E_1$  is Hurwitz, Lemma 4 in Appendix A.5 applied with  $k = 1$  and  $X = \mathcal{E}_1$ , can be used to claim that there exist a  $\bar{\nu}_1 > 0$  and, for any  $T_1 > 0$ , a  $\underline{\ell}_1 \geq 1$  such that for all  $\ell \geq \underline{\ell}_1$  and all  $\nu(t)$  satisfying  $\|\nu(\cdot)\|_\infty < \bar{\nu}_1/\ell$  the following holds

$$|\ell \tilde{\chi}_1(t)| \leq 1 \quad \forall t \geq T_1.$$

Hence, by noting that

$$\begin{aligned}\ell^{-1} |\eta_1 - x_2| &\leq |\tilde{\chi}_{12}| \leq |\tilde{\chi}_1| \\ |\eta_1| &\leq |\eta_1 - x_2| + |x_2|\end{aligned}$$

holds for any  $\ell \geq 1$ , we get

$$|\eta_1(t)| \leq \ell |\tilde{\chi}_1(t)| + |x_2(t)| \leq r_2 + 1$$

for any  $\ell \geq \underline{\ell}_1$  and for all  $t \geq T_1$ , namely  $\text{sat}_{r_2}(\eta_1(t)) = \eta_1(t)$  for all  $t \geq T_1$ .

**General case  $i \leq n-1$ .** Proceeding by induction, we assume that there exist a  $T_{i-1} > 0$ , a  $\bar{\nu}_{i-1} > 0$  and  $\underline{\ell}_{i-1} > 0$  such that for all  $\ell \geq \underline{\ell}_{i-1}$  and all  $\nu(t)$  satisfying  $\|\nu(\cdot)\|_\infty < \bar{\nu}_{i-1}/\ell^{i-1}$  then  $\text{sat}_{r_{j+1}}(\eta_j(t)) = \eta_j(t)$  for all  $j = 1, \dots, i-1$  and  $t \geq T_{i-1}$ . By using the change of coordinates (A.2) and using the notation  $\tilde{\chi}_{[k]} = (\tilde{\chi}_1, \dots, \tilde{\chi}_k)^T$  for the first  $k$ -th components of  $\tilde{\chi}$ , it turns out that for  $t \geq T_{i-1}$

$$\begin{aligned}\begin{pmatrix} \dot{\tilde{\chi}}_{[i-1]} \\ \dot{\tilde{\chi}}_i \end{pmatrix} &= \ell \begin{pmatrix} M_{i-1} & 0 \\ K_i B_{2(i-1)}^T & E_i \end{pmatrix} \begin{pmatrix} \tilde{\chi}_{[i-1]} \\ \tilde{\chi}_i \end{pmatrix} + \ell \bar{K}_1 \nu(t) \\ &\quad + \frac{1}{\ell^i} \begin{pmatrix} B_{2(i-1)} & 0 \\ 0 & B_2 \end{pmatrix} \begin{pmatrix} \ell u_{i-1}(t) \\ u_i(t) \end{pmatrix}\end{aligned}$$

where  $\bar{K}_1 := \text{col}(K_1, 0, \dots, 0)$ . Note that

$$\begin{aligned}u_{i-1}(t) &= \text{sat}_{r_{i+1}}(\eta_i) - x_{i+1} \\ &= \text{sat}_{r_{i+1}}(\ell^i B_2^T \tilde{\chi}_i + x_{i+1}) - x_{i+1}.\end{aligned}$$

By the Lipschitz mean-value theorem, there exists a  $s(\cdot) \in \mathcal{C}_{[0,1]}$  such that

$$u_{i-1}(t) = s(t) \ell^i B_2^T \tilde{\chi}_i$$

from which the  $\tilde{\chi}_{[i]}$  dynamics can be written as

$$\dot{\tilde{\chi}}_{[i]} = \ell \Lambda_i(s(t)) \tilde{\chi}_{[i]} + \frac{1}{\ell^i} B_{2i} u_i(t) + \ell \bar{K}_i \nu(t)$$

where  $\Lambda_i(s(t))$  is defined as in (13). Note that the initial condition  $\varepsilon_{[i]}(0)$  ranges in a compact set  $\mathcal{E}_i$  not dependent on  $\ell$  (for all  $\ell \geq 1$ ). By Lemma 4 in Appendix A.5 applied with  $k = i$  and  $X = \mathcal{E}_i$ , it turns out that there exist a  $\bar{\nu}_i \leq \bar{\nu}_{i-1}$  and, for all  $T_i > T_{i-1}$ , a  $\underline{\ell}_i \geq \underline{\ell}_{i-1}$  such that for all  $\ell \geq \underline{\ell}_i$  and all  $\nu(t)$  satisfying  $\|\nu(\cdot)\|_\infty < \bar{\nu}_i / \ell^i$  the following holds

$$|\ell^i \tilde{\chi}_{[i]}(t)| \leq 1 \quad \forall t \geq T_i.$$

From this, by noting that (for any  $\ell \geq 1$ )

$$\ell^{-i} |\eta_i - x_{i+1}| \leq |\tilde{\chi}_{i2}| \leq |\tilde{\chi}_{[i]}|$$

and

$$\begin{aligned} |\eta_i| &\leq |\eta_i - x_{i+1}| + |x_{i+1}| \leq |\ell^i \tilde{\chi}_{i2}| + |x_{i+1}| \\ &\leq |\ell^i \tilde{\chi}_{[i]}| + |x_{i+1}| \end{aligned}$$

it follows that for all  $\ell \geq \underline{\ell}_i$  and any  $\nu(t)$  fulfilling  $\|\nu(\cdot)\|_\infty < \bar{\nu}_i / \ell^i$

$$|\eta_i(t)| \leq \ell^i |\tilde{\chi}_{[i]}| + |x_{i+1}| \leq r_{i+1} + 1 \quad \forall t \geq T_i,$$

namely  $\text{sat}_{r_{i+1}}(\eta_i(t)) = \eta_i(t)$  for all  $t \geq T_i$ . This proves item (b) of the proposition.  $\square$

### A.5 Auxiliary Lemmas

**Lemma 3** *Let consider the system*

$$\dot{x} = \ell A(s(t))x + \frac{1}{\ell^k} \Delta(x, d) + \ell K \nu(t)$$

where  $s \in \mathcal{C}_{[0,1]}$ , with state  $x \in \mathbb{R}^n$ , bounded disturbances  $d$  and  $\nu$ , and with  $k$  and  $\ell$  positive numbers. Assume the following:

i) *there exists a  $P = P^T > 0$  such that for all  $s \in [0, 1]$  the following holds*

$$PA(s) + A(s)^T P \leq -I;$$

i)  $\Delta(x, d) \leq L|x| + R$  for some  $L > 0$ ,  $R > 0$ .

*Then, there exist positive constants  $\mu_i$ ,  $i = 1, \dots, 4$  and  $\underline{\ell} \geq 1$  such that for all  $\ell \geq \underline{\ell}$  and for all  $s(\cdot) \in \mathcal{C}_{[0,1]}$*

$$|x(t)| \leq \mu_1 \exp(-\ell \mu_2 t) |x(0)| + \frac{\mu_3}{\ell^{k+1}} \|d(\cdot)\|_\infty + \mu_4 \|\nu(\cdot)\|_\infty.$$

**Proof:** The proof follows by elementary Lyapunov arguments that are omitted.  $\square$

The following lemma immediately comes from Lemma 3 by picking  $L = 0$ .

**Lemma 4** *Consider the system*

$$\dot{x} = \ell A(s(t))x + \frac{1}{\ell^k} B u(t) + \ell K \nu(t) \quad (\text{A.16})$$

where  $s \in \mathcal{C}_{[0,1]}$ , with state  $x \in \mathbb{R}^n$ , and with  $k$  and  $\ell$  positive numbers. Assume there exists a  $\bar{u} > 0$  such that  $\|u(\cdot)\|_\infty < \bar{u}$  and that  $A(s)$  satisfies the assumption in item (i) in the previous lemma.

Then, there exists  $\bar{\nu} > 0$  and, for any compact set  $X \subset \mathbb{R}^n$  and  $T > 0$ , there exists  $\underline{\ell} \geq 1$  such that for any  $\ell \geq \underline{\ell}$  and any  $s(\cdot) \in \mathcal{C}_{[0,1]}$  trajectories of (A.16) originating from  $X$  and subject to disturbances  $\nu(t)$  fulfilling  $\|\nu(\cdot)\|_\infty \leq \bar{\nu} / \ell^k$ , satisfy

$$|\ell^k x(t)| \leq 1 \quad \forall t \geq T.$$

## References

- [1] J. H. Ahrens and H. K. Khalil, “High-gain observer in the presence of measurement noise: a switched-gain approach”, *Automatica*, vol. 45, pp. 936–943, 2009.
- [2] D. Astolfi, “Observers and Robust Output Regulation for Nonlinear Systems”, *Ph.D. Thesis* with University of Bologna and PLS Mines ParisTech, 2016.
- [3] D. Astolfi and L. Marconi, “A high-gain nonlinear observer with limited gain power”, *IEEE Transaction on Automatic Controls*, vol. 60, no. 11, pp. 3059–3064, 2015.
- [4] D. Astolfi, L. Marconi, L. Praly and A. Teel “Sensitivity to High-Frequency Measurement Noise of Nonlinear High-Gain Observers”, *10th IFAC Symposium on Nonlinear Control Systems*, 2016
- [5] D. Astolfi, L. Marconi and A. Teel “Low-power peaking-free high-gain observers for nonlinear systems”, *15th European Control Conference*, 2016.
- [6] D. Astolfi and L. Praly, “Output feedback stabilization for SISO nonlinear systems with an observer in the original coordinates”, *IEEE 52nd Conference on Decision and Control*, pp 5927–5932, 2013.
- [7] D. Astolfi, L. Praly and L. Marconi, “A note on observability canonical forms for nonlinear Systems”, *9th IFAC Symposium on Nonlinear Control Systems*, pp 436–438, 2013.
- [8] A. N. Atassi and H. K. Khalil, “A separation principle for the stabilization of a class of nonlinear systems”, *IEEE Transactions on Automatic Control*, vol. 44, pp 1672–1687, 1999.

- [9] G. Besançon, *Nonlinear observers and applications*, Springer, 2007.
- [10] M. Bin, D. Astolfi and L. Marconi, “Robust Internal Model Design by Nonlinear Regression via Low-Power High-Gain Observers”, to appear, *55th IEEE Conference on Decision and Control*, 2016.
- [11] J. 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.
- [12] J. P. Gauthier and G. Bornard, “Observability for any  $u(t)$  of a class of nonlinear systems”, *IEEE Transactions on Automatic Control*, vol. 26, no. 4, pp. 922–926, 1981.
- [13] J.P. Gauthier and I. Kupka, “*Deterministic observation theory and applications*”, Cambridge University Press, 2004.
- [14] H. K. Khalil and L. Praly, “High-gain observers in nonlinear feedback control”, *Int. Journal of Robust and Nonlinear Control*, vol. 24, pp 993–1015, 2014.
- [15] C. I. Byrnes and A. Isidori, “Nonlinear Internal Models for Output Regulation”, *IEEE Trans. Automatic Control*, vol. 49, no. 12, pp. 2244–2247, 2004.
- [16] A. Isidori and C.I. Byrnes, “Steady-State Behaviors in Nonlinear Systems with an Application To Robust Disturbance Rejection”, *Annual Reviews in Control*, vol. 32, pp. 1–16, 2008.
- [17] P. Lancaster, L. Rodman, *Algebraic Riccati Equations*, Clarendon Press, Oxford, 1995.
- [18] M. Maggiore, K.M. Passino, “A separation principle for a class of non uniformly completely observable systems”, *IEEE Transactions on Automatic Control*, vol. 48, no. 7, pp. 1122–1133, 2003.
- [19] R. Marino and G. L. Santosuosso, “Nonlinear Observers of Time Derivatives from Noisy Measurements of Periodic Signals” in *Nonlinear control in the year 2000*, vol. 2, pp. 123–135, Springer-Verlag, 2001.
- [20] R. Martinez-Guerra and J. L. Mata-Machuca, *Fault Detection and Diagnosis in Nonlinear Systems*, Springer Verlag, 2014.
- [21] M.S. Nazrulla and H.K. Khalil, “Robust stabilization of nonminimum phase nonlinear systems using extended high gain observers”, *IEEE Transactions on Automatic Control*, vol. 56, 802–813, 2011.
- [22] A. A. Prasov and H. K. Khalil, “A Nonlinear High-Gain Observer for Systems With Measurement Noise in a Feedback Control Framework”, *IEEE Transactions on Automatic Controls*, vol. 58, no. 3, pp. 569–580, 2013.
- [23] C. Prieur, S. Tarbouriech and L. Zaccarian “Hybrid high-gain observers without peaking for planar nonlinear systems”, *IEEE 51st Conference on Decision and Control*, pp. 6175–6180, 2012.
- [24] W. J. Rugh, *Linear System Theory*, Prentice-Hall, Englewood Cliffs, 1996.
- [25] R. Sanfelice and L. Praly, “On the performance of high-gain observers with gain adaptation under measurement noise”, *System & Control Letters*, vol. 47, pp. 2165–2176, 2011.
- [26] A. Teel, “Further variants of the Astolfi/Marconi high-gain observer”, *American Control Conference*, pp. 993-998, 2016.
- [27] A. Teel and L. Praly, “Global stabilizability and observability imply semi-global stabilizability by output feedback”, *Systems & Control Letters*, vol. 22, pp. 313–325, 1994.
- [28] A. Teel and L. Praly, “Tools for semiglobal stablization by partial state and output feedback”, *SIAM J. Control Optim.*, vol. 33, no. 5, pp 1443–1488, 1995.
- [29] L. Wang, D. Astolfi, H. Su, and L. Marconi, “High-gain observers with limited gain power for systems with observability canonical form”, accepted to *Automatica*.
- [30] L. Wang, D. Astolfi, S. Hongye, L. Marconi and A. Isidori, “Output Stabilization for a Class of Nonlinear Systems Via High-Gain Observer with Limited Gain Power”, *IFAC 1st Conference on Modelling, Identification and Control of Nonlinear Systems*, vol. 48, no. 11, pp. 730–735, 2015.