Low-Power Peaking-Free High-Gain Observers

Daniele Astolfi [a], Lorenzo Marconi [b], Laurent Praly [c], Andrew Teel [d]

[a] *Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France*

[b] *CASY - DEI, University of Bologna, Italy*

[c] *MINES ParisTech, PSL Research University, CAS - Centre automatique et systèmes, Paris, France*

[d] *Electrical and Computer Engineering Department, University of California, Santa Barbara, CA 93106-9560 USA*

Abstract

We propose a *peaking-free low-power* high-gain observer that preserves the main feature of standard high-gain observers in terms of arbitrarily fast converge to zero of the estimation error, while overtaking their main drawbacks, namely the "peaking phenomenon" during the transient and the numerical implementation issue deriving from the high-gain parameter that is powered up to the order of the system. Moreover, the new observer is proved to have superior features in terms of sensitivity of the estimation error to high-frequency measurement noise when compared with standard high-gain observers. The proposed observer structure has an high-gain parameter that is powered just up to two regardless the dimension of the observed system and adopts saturations to prevent the peaking of the estimates during the transient. As for the classical solution, the new observer is robust with respect to uncertainties in the observed system dynamics in the sense that practical estimation in the high-gain parameter can be proved.

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 Khalil and Praly (2014) and references therein). See also their use in the separation principles Atassi and Khalil (2000), output feedback stabilization Teel and Praly (1994), output regulation Byrnes and Isidori (2004) or fault detection Martinez-Guerra and Mata-Machuca (2013).

In the design of a "standard" high-gain observer, the high-gain parameter, denoted as ℓ 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 shows large peaks of a magnitude that are proportional to ℓ^{n-1}. Last but not least, high-gain observers are known for having high-sensitivity to high-frequency measurement noise, which makes state estimates practically unusable especially when the dimension n is very large. In order to address the peaking phenomenon, different schemes have been proposed in Astolfi and Praly (2017) and Maggiore and Passino (2003). In Astolfi and Praly (2017), the authors modify the observer dynamics under a convexity assumption in order to constrain the state of the observer in some prescribed convex closed set. This technique can be applied to multi-input multi-output nonlinear systems. In Maggiore and Passino (2003), the authors deal with peaking by means of a projection approach. In order to improve the sensitivity to measurement noise, the majority of researchers focused on schemes with time-varying gains, either with switched approaches, Ahrens and Khalil (2009), or with adaptive design, Boizot et al. (2010), Sanfelice and Praly (2011). Recently, in Khalil and Priess (2016), a low-pass filter has been proposed

⋆ Research supported in part by the "Région Grand-Est" of France, by the European Project SHERPA (ICT 600958), by AFOSR grant FA9550-15-1-0155 and NSF grant ECCS-1508757.

Email addresses: astolfi.daniele@univ-lorraine.fr (Daniele Astolfi), lorenzo.marconi@unibo.it (Lorenzo Marconi), Laurent.Praly@mines-paristech.fr (Laurent Praly), teel@ece.ucsb.edu (Andrew Teel).

in order to reduced the effect of measurement noise in output feedback stabilization problems.

A new high-gain observer able to overtake some of the drawbacks of classical structures has been recently proposed in Astolfi and Marconi (2015). 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 ℓ and ℓ^2 while preserving the same behaviours of a standard high-gain observer. The new construction relies on an interconnected cascade of $n-1$ high-gain observers of dimension two. This observer practically solves the aforementioned challenging problem of numerical implementation present in standard high-gain observers. 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 Astolfi and Marconi (2015) 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 Wang et al. (2017). 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 and output regulation, Astolfi et al. (2017). Although the new observer structure solves the problem of numerical implementation, the peaking phenomenon is still present. This has been partially solved in Astolfi, Marconi and Teel (2016), by adding saturations at various levels in the observer structure. With the proposed technique, it is possible to remove the peaking from the first $n-1$ state estimates. Along this route, two similar schemes, which follow the seminal idea presented in Astolfi and Marconi (2015), have been recently proposed, in Teel (2016) and Khalil (2017), to address the implementation issues and the peaking phenomenon. In Teel (2016), the author shows how to build a high-gain observer by interconnecting a cascade of reduced order high-gain observer of dimension 1. A simpler scheme, without feedback interconnection terms, that can not ensure asymptotic estimate, is presented in Khalil (2017). It is worth stressing, however, that even if the dimension of the observers is n, neither scheme improves the sensitivity properties with respect to standard high-gain observers.

The objective of this work is twofold. On the one hand, we combine the recent ideas of Astolfi and Marconi (2015) and Astolfi, Marconi and Teel (2016) to propose an observer of dimension $2n-1$ which is still "low power" (namely it uses only gains proportional to ℓ and ℓ^2) and yet eliminates the peaking phenomenon. This is achieved by appropriately adding saturation functions in the observer dynamics. In particular, the n estimates provided by the proposed observer are peaking-free while the additional $n-1$ auxiliary variables may reach values proportional to ℓ (and not to ℓ^{n-1} as in standard high-gain observers) during the transient. The resulting gain choices and transient behaviours address the numerical challenge. On the other hand, we fully characterise the sensitivity to high-frequency measurement noise for nonlinear systems by showing the improvement with respect to standard high-gain observers. This is done by extending the analysis tool recently introduced in Astolfi, Marconi, Praly and Teel (2016) 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 Astolfi and Marconi (2015), but similar results hold for the wider class of systems in feedback form Wang et al. (2017).

The paper is organized as follows. We present the framework and we recall the high-gain observer technique in Section 2. Then, we provide the main results in Section 3. A simulation example is given in Section 4. The proofs of the main results are detailed in Section 5. Conclusions are discussed in Section 6. Some technical lemmas are given in Appendix A.

Notation. $\mathbb{R}$ denotes the field of real numbers and, for $x \in \mathbb{R}^n$, $|x|$ denotes the Euclidean norm of x. With $s : \mathbb{R}_{\geq 0} \to \mathbb{R}^m$ a bounded signal, we define $\|s\|_a^b := \sup_{t\in[a,b)} |s|$ and $\|s\|_\infty := \|s\|_0^\infty$. 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

$$A_i = \begin{pmatrix} 0_{i-1,1} & I_{i-1} \\ 0 & 0_{1,i-1} \end{pmatrix}, \; B_i = \begin{pmatrix} 0_{i-1,1} \\ 1 \end{pmatrix}, \; C_i^T = \begin{pmatrix} 1 \\ 0_{i-1,1} \end{pmatrix},$$

where $0_{i,j}$ denotes a matrix of dimension $i \times j$ containing zeros everywhere, and I_i denotes the identity matrix of dimension i. For $r > 0$, a saturation function $\mathrm{sat}_r : \mathbb{R} \to \mathbb{R}$ is any strictly increasing C^1 function satisfying

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

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

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 Gauthier and Kupka (2001))

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

where $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ is the state, y is the measured output with ν an additive unknown measurement noise, and $t \mapsto d(t) \in \mathbb{R}^{n_d}$, $n_d > 0$, is any (unknown) bounded signal that may represent parametric uncertainties in the function $\varphi(\cdot, \cdot)$ or unknown disturbances. The following assumption is made throughout the paper.

Assumption 1 *The compact sets $D \subset \mathbb{R}^{n_d}$ and $X \subset \mathbb{R}^n$ and the positive $\bar{\varphi}_x > 0$ are such that*

- *$d(t) \in D$ and $x(t) \in X$ for all $t \geq 0$;*
- *$|\varphi(x_1, d) - \varphi(x_2, d)| \leq \bar{\varphi}_x |x_1 - x_2|$ for all $x_1, x_2 \in X$ and for all $d \in D$.*

We observe that all the forthcoming analysis could 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.

In the previous framework, we are interested in the semi-global high-gain observation problem, namely in the design of an asymptotic observer with a rate of convergence that can be made arbitrarily fast by tuning a single parameter (see Khalil and Praly (2014) and references therein).

The standard high-gain observer for the class of systems (1) is given by

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

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

$$e_1 := y - \hat{x}_1 , \tag{2b}$$

$k_1, \dots, k_n$ 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 $K := \mathrm{col}(k_1, \dots, k_n)$, the matrix $A_n - KC_n$ is Hurwitz, and taking the high-gain parameter ℓ large enough in relation to the Lipschitz constant of $\varphi(\cdot, \cdot)$ on $X \times D$. In particular, under Assumption 1, it is possible to prove that, by letting $\ell^\star := 2\, \bar{\varphi}_x\, |P|$, 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 $\ell \geq \ell^\star$ the estimation errors provided by the observer (2) satisfy the following bounds for all $t \geq 0$

$$\begin{aligned} |\hat{x}_i(t) - x_i(t)| \;\leq\; & c_1\, \ell^{i-1} \exp(-c_2\, \ell\, t) |\hat{x}(0) - x(0)| \\ & + \frac{c_3}{\ell^{n+1-i}} \, \|d\|_\infty + c_4 \ell^{i-1} \, \|\nu\|_\infty \end{aligned} \tag{3}$$

for $i = 1, \dots, n$, and for some positive constants c_i, $i = 1, \dots, 4$, independent of ℓ. This, in particular, can be easily established by making the change of coordinates

$$\hat{x} = \mathrm{col}(\hat{x}_1, \dots, \hat{x}_n) \to \chi = \mathrm{col}(\chi_1, \dots, \chi_n), \qquad \chi_i := (\hat{x}_i - x_i)/\ell^{i-1},$$

and by using the Lyapunov function $V = \chi^T P \chi$ (see also Lemma 4 in Appendix A or Khalil and Praly (2014) for a detailed proof). 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 ℓ showing up in the exponential function. This, in turn, implies that in nominal conditions (namely when there is no measurement noise and the disturbance d is constantly zero), the true value of the state variable can be practically recovered in an arbitrarily small amount of time and exponentially convergence of the estimate is guaranteed. The term proportional to ℓ^{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 errors assume large values in the initial observation time if ℓ 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 is that the observer (2) is input-to-state stable (ISS) with respect to the disturbance inputs d and ν. 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 ℓ. As for the measurement noise ν, on the other hand, the asymptotic gain increases proportionally to ℓ^{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, another limit of the high-gain structure (2) is the presence of ℓ 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 (3) refers to the so-called $\mathcal{L}_\infty$ gain, namely characterises 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 the low-pass filtering properties of the observer (2). This high-frequency characterisation of the asymptotic gain has been fully characterised in Astolfi, Marconi, Praly and Teel (2016) whose main result is briefly recalled hereafter. We consider, in particular, the measurement noise as a quasi-periodic signal of the form

$$\nu(t) \;=\; \sum_{i=1}^{n_\nu} \nu_i^c \cos\left(\frac{\omega_i}{\varepsilon}\, t\right) + \nu_i^s \sin\left(\frac{\omega_i}{\varepsilon}\, t\right) \qquad (4)$$

where n_ν, ν_i^c, ν_i^s, ω_i are positive numbers and where $\varepsilon \in (0,1)$ is a small number parametrising the frequencies of the signal $\nu(t)$. The main result proved in Astolfi, Marconi, Praly and Teel (2016) is the following.

Proposition 1 *Consider system* (1)*,* (2) *and suppose that Assumption 1 holds, $d(t) \equiv 0$ for all $t \geq 0$, and ν is generated by* (4)*. Let $\ell > 1$ be fixed so the bound* (3) *holds. Then, there exist $\varepsilon^\star(\ell) > 0$ and $\hat{c} > 0$ such that, for all positive $\varepsilon \leq \varepsilon^\star(\ell)$, the following holds*

$$\limsup_{t\to\infty} |\hat{x}_i(t) - x_i(t)| \;\leq\; \varepsilon\, \hat{c}\, \ell^i\, \|\nu\|_\infty \quad i = 1,\ldots,n.$$

Proposition 1 shows that, once ℓ is fixed, the sensitivity of the estimation error to measurement noise decreases as ε takes smaller values, namely as higher frequency noise signals are considered, with an asymptotic gain proportional to ε. For linear systems, this property immediately comes by frequency response arguments using the fact that the relative degree between the measurement noise ν and the estimation error $x_i - x_i$ for (1)-(2) is unitary for all $i = 1,\ldots,n$. The extension to nonlinear systems of the form (1) is more involved and can be found in Astolfi, Marconi, Praly and Teel (2016). It is worth noting that the analysis in Astolfi, Marconi, Praly and Teel (2016) is based on the assumption that the measurement noise is generated as in (4), while, in practice, measurement noise is usually white or coloured random noise. However, simulations confirm that the result of Proposition 1 provides a good indication of the true performance of the observer in presence of coloured noise. This will be discussed further in Section 4.

3 Main Results

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 ℓ 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 Astolfi and Marconi (2015) that has dimension $2n-2$. The motivation for extending the state of the observer by one with respect to the solution provided in Astolfi and Marconi (2015) is to pave the way for the "peaking-free" solution presented in Section 3.3.

The structure of the proposed observer is composed of n blocks, where each of the first $n-1$ blocks has dimension 2 and the last one has dimension 1. The two state components of the i-th block for $i = 1,\ldots,n-1$ 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 ($i = 1,\ldots n-1$) 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}), ℓ is the high gain parameter and (α_i, β_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 α_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 η_{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 in this way (see Figure 1), we obtain the proposed observer

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

where $\hat{x} = \mathrm{col}(\hat{x}_1,\ldots,\hat{x}_n) \in \mathbb{R}^n$, $\eta = \mathrm{col}(\eta_1,\ldots,\eta_{n-1}) \in \mathbb{R}^{n-1}$ is the state, $\underline{\alpha} := \mathrm{col}\,(\alpha_1\,,\,\ldots\,,\,\alpha_n) \in \mathbb{R}^n$ and $\underline{\beta} := \mathrm{col}\,(\beta_1\,,\,\ldots\,,\,\beta_{n-1}) \in \mathbb{R}^{n-1}$ are design parameters and ℓ the high-gain parameter, and the variables e_i, $i = 1,\ldots,n$ are defined as

$$\begin{aligned} e_1 &:= y - \hat{x}_1 \\ e_i &:= \eta_{i-1} - \hat{x}_i, \qquad i = 2,\ldots,n\,. \end{aligned} \qquad (5b)$$

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 $K_i := \mathrm{col}(\alpha_i, \beta_i)$ and $E_i := A_2 - K_iC_2$, $i = 1,\ldots,n-1$, let the matrices $M_i \in \mathbb{R}^{2i\times 2i}$, $i =$

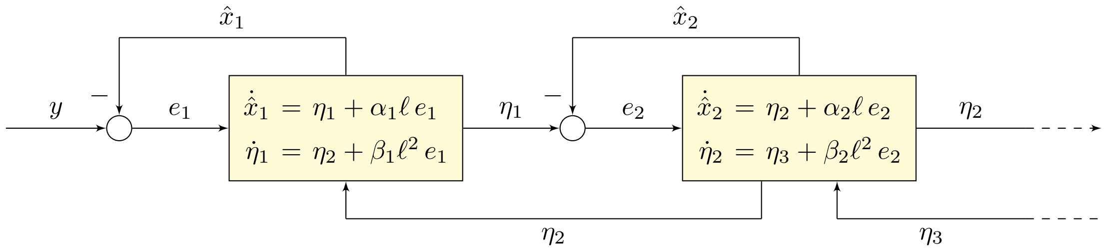

Fig. 1. Block diagram representation of the low-power high-gain observer (5).

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

$$M_i := \begin{pmatrix} M_{i-1} & B_{2(i-1)} B_2^T \\ K_i B_{2(i-1)}^T & E_i \end{pmatrix}, \quad i = 2, \dots, n-1,$$
$$M_n := \begin{pmatrix} M_{n-1} & 0 \\ \alpha_n B_{2(n-1)}^T & -\alpha_n \end{pmatrix}.$$

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*

$$P M_n + M_n^T P = -I\,. \tag{6}$$

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 Section 3.4). With the matrix M_n Hurwitz and the high-gain parameter ℓ taken sufficiently large, the estimation error $\hat{x} - x$ provided by the observer (5) can be shown to fulfil the same bound (3) yielded by the standard high-gain observer (2). This is detailed in the next theorem in which we define $\boldsymbol{x} \in \mathbb{R}^{2n-1}$ and $\hat{\boldsymbol{x}} \in \mathbb{R}^{2n-1}$ as

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

Theorem 1 *Consider system* (1) *under the Assumption 1. Consider the observer* (5) *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* (6) *fulfilled for some $P = P^T > 0$. Furthermore, let $\ell^\star := 2\,\bar{\varphi}_x\,|P|$. Then there exist $\mu_i > 0$, $i = 1, \dots, 4$, such that for any $\ell > \ell^\star$ the following bounds hold*

$$|\hat{x}_i(t) - x_i(t)| \;\leq\; \ell^{i-1}\mu_1 \,\exp(-\ell\,\mu_2 t)\,|\hat{\boldsymbol{x}}(0) - \boldsymbol{x}(0)| + \frac{\mu_3}{\ell^{n+1-i}}\|d\|_\infty + \mu_4\,\ell^{i-1}\|\nu\|_\infty \tag{7}$$

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

$$|\eta_i(t) - x_{i+1}(t)| \;\leq\; \ell^{i}\mu_1 \,\exp(-\ell\,\mu_2 t)\,|\hat{\boldsymbol{x}}(0) - \boldsymbol{x}(0)| + \frac{\mu_3}{\ell^{n-i}}\|d\|_\infty + \mu_4\,\ell^{i}\|\nu\|_\infty \tag{8}$$

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 all $t \geq 0$.

The proof of this theorem is deferred to Section 5.1. 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 $(\eta_1, \dots, \eta_{n-1})$. Furthermore, we observe that the lower bound $\ell^\star$ 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 $\ell^\star$ might be different. As clear from the bounds (7)-(8), 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 in 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 ℓ 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 ν and the estimation error in terms of the parameter ε. 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) is that the relative degree between the "input" ν 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

$$m := \left\lceil \frac{n+1}{2} \right\rceil \tag{9}$$

and considering a general case in which the function $\varphi(x)$ is affected by x_1, the relative degree in question is i for $i = 1, \ldots, m$ and $n - i + 2$ for $i = m+1, \ldots, n$. This property is at the basis of the next proposition whose proof is presented in Section 5.2.

Proposition 2 *Consider system* (1), (5) *and suppose that Assumption 1 holds,* $d(t) \equiv 0$ *for all* $t \geq 0$, *and* ν *is generated by* (4). *Let* $\underline{\alpha} \in \mathbb{R}^n$ *and* $\underline{\beta} \in \mathbb{R}^{n-1}$ *and* $\ell > 1$ *be fixed according to the statement of Theorem 1. Then, there exist* $\varepsilon^\star(\ell) > 0$ *and* $\hat{c} > 0$ *such that, for all positive* $\varepsilon \leq \varepsilon^\star(\ell)$, *the following holds*

$$\limsup_{t \to \infty} |\hat{x}_i(t) - x_i(t)| \leq \varepsilon^i \, \hat{c} \, \ell^{2i-1} \, \|\nu\|_\infty$$

for $i = 1, \ldots, m$, *and*

$$\limsup_{t \to \infty} |\hat{x}_i(t) - x_i(t)| \leq \varepsilon^{n-i+2} \, \hat{c} \, \ell \, \|\nu\|_\infty$$

for $i = m+1, \ldots, n$.

Proposition 2 shows that observer (5) behaves as "low-pass" filter, namely

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

In addition, the remarkable feature of observer (5) is to have an asymptotic gain between the measurement noise and the i-th error component that is proportional to ε 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 ε 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 ε tends to zero, namely as far as high-frequency noise is concerned. The numerical analysis in Section 4 will provide further insights on the benefits of the new observer over the standard one.

3.3 Peaking-free low-power observer

In this section we show how the observer (5) 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 the saturation function given in the Notation and by defining $r_i > 0$ as

$$r_i := \max_{x \in X} |x_i| \quad i = 1, \ldots, n \, , \tag{10}$$

the low-power peaking-free observer (see Figure 2) takes the form (compare with (5))

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

with

$$\begin{aligned}
e_1 &:= y - \hat{x}_1 \\
e_i &:= \mathrm{sat}_{r_i}(\eta_{i-1}) - \hat{x}_i, \qquad i = 2, \ldots, n \, ,
\end{aligned} \tag{11b}$$

where $\hat{x} = \mathrm{col}(\hat{x}_1, \ldots, \hat{x}_n) \in \mathbb{R}^n$, $\eta = \mathrm{col}(\eta_1, \ldots, \eta_{n-1}) \in \mathbb{R}^{n-1}$ is the state, $\underline{\alpha} := \mathrm{col}\,(\alpha_1 \, , \, \ldots \, , \, \alpha_n) \in \mathbb{R}^n$ and $\underline{\beta} := \mathrm{col}\,(\beta_1 \, , \, \ldots \, , \, \beta_{n-1}) \in \mathbb{R}^{n-1}$ are positive coefficients to be properly chosen, and ℓ 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 section, let $\Lambda_i : [0,1] \to \mathbb{R}^{2i \times 2i}$, $i = 1, \ldots, n-1$, $\Lambda_n \in \mathbb{R}^{2n-1 \times 2n-1}$ be continuous matrices defined as $\Lambda_1 := M_1$,

$$\Lambda_i(s) := \begin{pmatrix} M_{i-1} & s \, B_{2(i-1)} \, B_2^T \\ K_i \, B_{2(i-1)}^T & E_i \end{pmatrix} \quad i = 2, \ldots, n-1 \tag{12}$$

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:*

- α_i, $i = 1, \ldots, n$ *and* β_i, $i = 1, \ldots, n-1$, *are all positive;*
- *for all* $i = 1, \ldots, n$, *there exist* $P_i = P_i^T > 0$ *and* $\mu_i > 0$ *such that for all* $s \in [0,1]$ *the resulting* $\Lambda_i(s)$ *fulfils*

$$P_i \Lambda_i(s) + \Lambda_i(s)^T P_i \leq -\mu_i I \, . \tag{13}$$

It turns out that the previous requirement can be always fulfilled by an appropriate choice of $\underline{\alpha}$ and $\underline{\beta}$. In particular, given a set of $\underline{\alpha}$ and $\underline{\beta}$ satisfying the "Low-power stability requirement" (see Definition 1), one may always check if the "Low-power strong stability requirement" is fulfilled by applying Lemma 2 in Section 3.4. Alternatively, one may design the coefficients $\underline{\alpha}$ and $\underline{\beta}$ by following the constructive procedure presented at the end of Section 3.4.

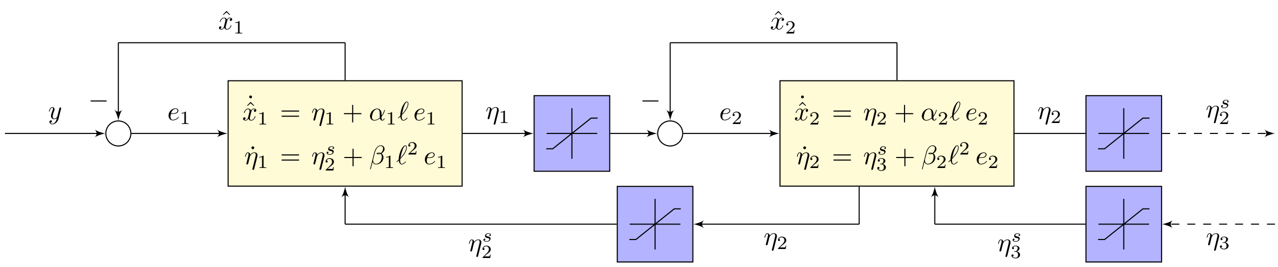

Fig. 2. Block diagram representation of the peaking free low-power high-gain observer (11). We denote $\eta_i^s = \mathrm{sat}_{i+1}(\eta_i)$ for the sake of compactness.

Proposition 3 *Consider system* (1) *under the Assumption 1. Consider the observer* (11) *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 $\bar{p}_i > 0$, $i = 2, \ldots, n$, *and, for each* $\bar{\nu} > 0$, *there exists* $\bar{p}_1 > 0$ *such that*

$$\begin{aligned} |\hat{x}_i(t) - x_i(t)| &\leq \bar{p}_i\,, \quad i = \ldots, n \\ |\eta_i(t) - x_{i+1}(t)| &\leq \ell\, \bar{p}_i\,, \quad i = 1, \ldots, n-1 \end{aligned} \tag{14}$$

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

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

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

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

The proof of the previous proposition is deferred to Section 5.3. Proposition 3 has two main consequences. First of all, the first inequality of (14), clearly shows that the estimates $\hat{x}_i$, $i = 1, \ldots, n$, given by the observer (11), do not peak with ℓ. In particular the ultimate bound $\|\hat{x}_i\|_\infty$ depends on the compact sets X, $\hat{X}$ and E but it is independent of ℓ. In other words, the observer (11) provides a peaking-free estimate $\hat{x}(t)$ of the state $x(t)$ of (1). On the other hand, the second inequality of (14) shows that the variables η_i may grow with ℓ during the transients, namely the value $\|\eta_i\|_\infty$ is proportional to the high-gain parameter ℓ. Nevertheless, it is worth noting that, from a computational point of view, this should not worry since the implementation of values proportional to ℓ is needed in order to design the gains of the observer. In other words, the maximum values that the auxiliary variables η_i may reach is of the same order of magnitude of the gains of the observers, allowing to implement numbers which are in general well-conditioned. This interesting feature is in general not guaranteed with other constructions, such as see Teel (2016).

The second consequence of the Proposition 3 is that the variables η_i, $i = 1 \ldots, n-1$, exit from saturation (item (b)) if the amplitude of the sensor noise is sufficiently small in relation to ℓ^i. In particular, if $\|\nu\|_\infty \leq \bar{\nu}/\ell^{n-1}$, the observer (11) boils down, in finite time, to the low-power observer (5) by thus recovering all the asymptotic properties detailed in Theorem 1 and Proposition 2. In some situations, though, the bounds in item (b) may be conservative (for instance when high-frequency measurement noise is considered). Finally, note that, to have item (b) fulfilled, the saturations level needs to be chosen large enough according to (10).

3.4 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 Astolfi and Marconi (2015). The MATLAB code for the design of the coefficients can be also found in Astolfi (2016).

Lemma 1 *Let* $\mathcal{P}(\lambda) = \lambda^{2n-1} + m_1\lambda^{2n-2} + \ldots + m_{2n-2}\lambda + m_{2n-1}$ *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 Astolfi and Marconi (2015), it turns out that the coefficients $(\alpha_1, \ldots, \alpha_{n-1})$ and $(\beta_1, \ldots, \beta_{n-1})$ can be designed to assign an arbitrary polynomial $\mathcal{P}_{n-1}(\lambda)$. From this the result immediately follows. $\square$

Given a set of coefficients $\underline{\alpha}$ and $\underline{\beta}$ one may check if also the "low-power strong stability requirement" in Definition 2 is fulfilled by direct application of the following Lemma.

Lemma 2 *Let $S_{i-1} = S_{i-1}^T > 0$, $S_i = S_i^T > 0$ and $\gamma_{i-1}, \gamma_i > 0$, be such that*

$$\begin{pmatrix} S_{i-1}M_{i-1} + M_{i-1}^T S_{i-1} & S_{i-1}B_{2(i-1)} & B_{2(i-1)} \\ B_{2(i-1)}^T S_{i-1} & -\gamma_{i-1} I & 0 \\ B_{2(i-1)}^T & 0 & -\gamma_{i-1} I \end{pmatrix} < 0, \tag{15}$$

$$\begin{pmatrix} S_i E_i + E_i^T S_i & S_i K_i & B_2 \\ K_i^T S_i & -\gamma_i I & 0 \\ B_2^T & 0 & -\gamma_i I \end{pmatrix} < 0\,. \tag{16}$$

If $\gamma_{i-1}\gamma_i < 1$, then there exists $P_i = P_i^T > 0$ and $\mu_i > 0$ such that (13) *holds.*

Proof: Consider the matrix $\Lambda_i(s)$ defined in (12) and regard it as the state matrix of a system resulting from the feedback interconnection of a first Hurwitz system

$$\begin{aligned} \dot{x}_{i-1} &= M_{i-1}x_{i-1} + B_{2(i-1)}v_{i-1} \\ y_{i-1} &= B_{2(i-1)}^T x_{i-1} \end{aligned} \tag{17}$$

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

$$\begin{aligned} \dot{x}_i &= E_i x_i + K_i v_i \\ y_i &= B_2^T x_i \end{aligned} \tag{18}$$

with state $x_i \in \mathbb{R}^2$, input $v_i \in \mathbb{R}$ and output $y_i \in \mathbb{R}$, under the feedback $v_{i-1} = s\, y_i$ and $v_i = y_{i-1}$. If M_i is Hurwitz, then, by applying the bounded real lemma see Lancaster and Rodman (1995), there exists S_{i-1} and γ_{i-1} such that inequality (15) holds and moreover the Lyapunov function $V_{i-1} = \gamma_{i-1} x_{i-1}^T S_{i-1} x_{i-1}$ satisfies

$$\dot{V}_{i-1} \le -\epsilon_{i-1}|x_{i-1}|^2 + \gamma_{i-1}^2|v_{i-1}|^2 - |y_{i-1}|^2$$

for some $\epsilon_{i-1} > 0$. Similarly, in view of (16), the Lyapunov function $V_i = \gamma_i x_i^T S_i x_i$ satisfies

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

for some $\epsilon_i > 0$. Now consider the composite Lyapunov function $W_i = V_{i-1} + aV_i$ where $a > 0$ is a real number to be selected. By using the previous inequalities and the definitions of v_{i-1}, v_i, we obtain

$$\begin{aligned} \dot{W}_i \le\; & -\epsilon_{i-1}|x_{i-1}|^2 + \gamma_{i-1}^2|v_{i-1}|^2 - |y_{i-1}|^2 \\ & \quad -a\epsilon_i|x_i|^2 + a\gamma_i^2|v_i|^2 - a|y_i|^2 \\ \le\; & -\epsilon_{i-1}|x_{i-1}|^2 - a\epsilon_i|x_i|^2 \\ & + \begin{pmatrix} y_{i-1}^T & y_i^T \end{pmatrix} \begin{pmatrix} (-1 + a\gamma_i^2)I & 0 \\ & (s^2\gamma_{i-1}^2 - a)I \end{pmatrix} \begin{pmatrix} y_{i-1} \\ y_i \end{pmatrix}. \end{aligned} \tag{19}$$

If $\gamma_i\gamma_{i-1} < 1$, then there exists $a > 0$ satisfying

$$\gamma_{i-1}^2 \le a \le \frac{1}{\gamma_i^2}$$

and therefore, for any $s \in [0,1]$, the inequality (19) reduces to $\dot{W}_i \le -\mu_i(|x_{i-1}|^2 + |x_i|^2)$ where $\mu_i = \min\{\epsilon_{i-1}, a\epsilon_i\}$. The matrix satisfying (13) is therefore $P_i := \mathrm{diag}(\gamma_{i-1}S_{i-1}, a\gamma_i S_i)$ with a satisfying the previous condition. This concludes the proof. $\square$

By recalling standard results on bounded real lemmas and equivalences between $\mathcal{L}_2$ and $\mathcal{H}_\infty$ gains, one may verify the condition $\gamma_{i-1}\gamma_i < 1$ by computing γ_{i-1}, γ_i as the $\mathcal{H}_\infty$ gains of the transfer functions of systems (17) and (18) instead of solving conditions (15) and (16). Moreover, by bearing in mind the definitions of E_i and K_i, we can compute the $\mathcal{H}_\infty$ gain of the transfer function of system (16) between input v_i and output y_i, which is β_i/α_i. As a consequence, the result of Lemma 2 can be directly applied to obtain a design procedure satisfying the "low-power strong stability requirement" as follows:

step 1) take (α_1, β_1) as any pair of positive numbers;
step i) for all $i = 2, \dots, n-1$, compute recursively α_i and β_i as any positive numbers such that $\frac{\beta_i}{\alpha_i} \ge \frac{1}{\gamma_{i-1}}$ with γ_{i-1} the $\mathcal{L}_2$ gain of system (17);
step n) take α_n as any positive number.

4 A numerical example

For illustration purposes we consider a system of the form (1) with $n = 5$ with the nonlinear function φ chosen as

$$\varphi(x) = 0.2(x_1^2 - 1) - x_2 - x_3 - 4x_4 - x_5\,. \tag{20}$$

As shown in Sprott (2010), system (1), (20) is a *crackle system* exhibiting chaotic behaviours, when the initial conditions are close enough to the origin, and possibly unstable otherwise. In the simulations we selected $x(0) = (-0.8, 0, 0, 0, 0)^T$. Numerical inspection shows that, with this initial conditions, $|x_1(t)| < 2.5$, $|x_2(t)| < 1$, $|x_i(t)| < 0.5$, $i = 3, 4, 5$, and $|\varphi(t)| \le 0.5$ for all $t \ge 0$. The observer (11) of dimension 9, has been implemented by following the prescriptions of Section 3. In particular, the coefficients $\underline{\alpha}$ and $\underline{\beta}$ have been chosen by following the recursive procedure of Astolfi and Marconi (2015) [1] obtaining $\alpha_i = 3$ for $i = 1, \dots, 4$, $\alpha_5 = 2$, $\beta_1 = 8.5714$, $\beta_2 = 3.2122$, $\beta_3 = 1.4267$, $\beta_4 = 0.5347$. In this way the poles of the matrix M_n are real and placed in the range $[-2, -1]$. It is also possible to verify by direct application of Lemma 2 that, with this choice,

[1] In particular we used the MATLAB code that can be found in Astolfi (2016).

the coefficients satisfy the low-power strong stability requirement of Definition 2. The saturation levels have been fixed to $r_i = 3$, $i = 1, \dots, 5$. The function $\varphi_s(\cdot)$ has been implemented by saturating the function φ such that $|\varphi_s(x)| \le 3$ for any $x \in \mathbb{R}^5$. Figure 1 shows the behaviour of the peaking-free low-power high-gain observer (11) when $\ell = 7$ with initial conditions chosen as $|\hat{x}_i(0)| = 1$, $i = 1, \dots, 5$, $|\eta_i(0)| = 0$, $i = 1, \dots, 4$, without measurement noise, namely $\nu(t) = 0$. Then, we compared the observer (11) with a standard high-gain observer (2) of dimension 5. The parameters k_i are chosen as $k_1 = 7.5$, $k_2 = 22.1875$, $k_3 = 32.3438$, $k_4 = 23.2188$, $k_5 = 6.5625$, so that the roots of the matrix $(A_5 - K_5 C_5)$, with $K_5 = (k_1, \dots, k_5)^T$, are real and equidistant in the range $[-2, -1]$. The function $\varphi_s(\cdot)$ in (2) is the same we used to implement observer (11).

In order to characterize the peaking phenomenon, we run numerous simulations with random initial conditions in the set $\{(x, \eta) \in \mathbb{R}^9 : |x_i| \le 3, i = 1, \dots, 5, |\eta_i| \le 3, i = 1, \dots, 4\}$. Table 1 shows, for different values of ℓ, the maximum peaking values of the state $(\hat{x}, \eta)$ of the low-power peaking free high-gain observer (11) and the time needed to converge to an error sufficiently small (*i.e.* the time T_ϵ such that $|x(t) - \hat{x}(t)| < \epsilon$ for all $t \ge T_\epsilon$) among all the simulations that we run. Table 2 shows, for different values of ℓ, the maximum peaking values of the state $\hat{x}$ of the high-gain observer (2) and the time T_ϵ. While the peaking on the estimates $\hat{x}_i$ provided by (2) is proportional to increasing powers of ℓ, as expected by the bound (3), no peaking is present on the estimates $\hat{x}_i$ provided by the peaking-free low-power high-gain observer (11). The peaking on the auxiliary state variables η_i is only proportional to ℓ, as expected by Proposition 3. Then, in order to characterize performance of the observer (11) in presence of measurement noise, we fixed the high-gain parameter as $\ell_1 = 7$. Note that the largest coefficient we need to implement in the observer (11) is in this case $\beta_1 \ell_1^2 = 420$. Similarly, we fixed $\ell_2 = 4.17$ for the high-gain observer (2) in order to practically match convergence rates of the two observers, namely to achieve the same T_ϵ with $\epsilon = 0.01$. In this case, note that the largest coefficient we need to implement is $k_5 \ell_2^5 = 8.27 \cdot 10^3$. We repeated the simulations in four different scenarios. In the the scenarios (a) and (c) we supposed that the measurement noise ν is some coloured noise generated by filtering white noise respectively with a band-pass filter (with band $[50 - 200]Hz$) and with a high-pass filter (with band $[1000 - \infty]Hz$). In contrast, in the scenarios (b) and (d) we supposed, as in Propositions 1 and 2, that the measurement noise ν is generated by a sinusoidal signal $\nu(t) = A \sin(\omega t)$. In the scenario (b) we considered $\omega = 100$ while in the scenario (d) we selected $\omega = 5000$. Tables 3 and 4 show the maximum value of the estimation errors $|\hat{x}_i(t) - x_i(t)|$, $i = 1, \dots, 5$, in steady-state for the two observers. The tables show the remarkable improvement in terms of disturbance attenuation (especially for the components $\hat{x}_i$, $i > 2$) of the low-power peaking free high-gain observer with respect to standard high-gain observer technique. For the observer (11) the measurement noise is attenuated on all components $\tilde{x}_i := \hat{x}_i - x_i$, $i = 1, \dots, 5$, while for the observer (2) the measurement noise is amplified for $i \ge 2$ when considering medium frequencies and for $i \ge 4$ at high-frequencies. Similar results are obtained when considering measurement noise with different amplitudes, frequencies or band-pass filters. Finally, we remark that the data collected in Tables 3, 4 confirm that the approximation given in Propositions 1 and 2 provides a good indicator of the steady-state behaviour of observers (2) and (11) in presence of coloured random measurement noise, thus supporting the validity of the proposed nonlinear analysis.

	$\ell = 7$	$\ell = 30$	$\ell = 70$
$T_{0.01}$	4.881	0.995	0.471
$\|\hat{x}_1\|_\infty$	3	3	3
$\|\hat{x}_2\|_\infty$	4.10	4.04	4.03
$\|\hat{x}_3\|_\infty$	3.69	3.61	3.59
$\|\hat{x}_4\|_\infty$	3.37	3.90	3.29
$\|\hat{x}_5\|_\infty$	3.20	3.05	3.02
$\|\eta_1\|_\infty$	43.89	181.81	421.98
$\|\eta_2\|_\infty$	33.36	135.15	312.22
$\|\eta_3\|_\infty$	18.43	69.93	159.90
$\|\eta_4\|_\infty$	9.19	30.80	68.35

Table 1
Converge time and peaking phenomenon of the low-power peaking-free high-gain observer (11) when $\nu(t) = 0$. Worst case for initial conditions in $x_i(0) = \{3, -3\}$ for $i = 1, \dots, 5$ and $\eta_i(0) = \{3, -3\}$ for $i = 1, \dots, 4$. $T_{0.01}$ is computed such that $|x(t) - \hat{x}(t)| < 0.01$ for all $t \ge T_{0.01}$.

	$\ell = 4.17$	$\ell = 20$	$\ell = 41.7$
$T_{0.01}$	4.901	1.136	0.604
$\|\hat{x}_1\|_\infty$	3	3	3
$\|\hat{x}_2\|_\infty$	41.84	194.15	403.42
$\|\hat{x}_3\|_\infty$	222.59	$4.92 \cdot 10^3$	$2.13 \cdot 10^4$
$\|\hat{x}_4\|_\infty$	613.59	$6.56 \cdot 10^4$	$5.91 \cdot 10^5$
$\|\hat{x}_5\|_\infty$	690.71	$3.55 \cdot 10^5$	$6.69 \cdot 10^6$

Table 2
Converge time and peaking phenomenon of the high-gain observer (2) when $\nu(t) = 0$. Worst case for initial conditions in $x_i(0) = \{3, -3\}$ for $i = 1, \dots, 5$. $T_{0.01}$ is computed such that $|x(t) - \hat{x}(t)| < 0.01$ for all $t \ge T_{0.01}$.

	(a)	(b)	(c)	(d)
$\|\nu\|_{\mathcal{T}_{0.01}}^{\infty}$	0.298	0.298	4.4894	4.4894
$\|\tilde{x}_1\|_{\mathcal{T}_{0.01}}^{\infty}$	0.059	0.065	0.017	0.0189
$\|\tilde{x}_2\|_{\mathcal{T}_{0.01}}^{\infty}$	0.280	0.263	0.004	0.0016
$\|\tilde{x}_3\|_{\mathcal{T}_{0.01}}^{\infty}$	0.660	0.404	0.0008	0.0004
$\|\tilde{x}_4\|_{\mathcal{T}_{0.01}}^{\infty}$	0.752	0.278	0.002	0.0020
$\|\tilde{x}_5\|_{\mathcal{T}_{0.01}}^{\infty}$	0.201	0.042	0.004	0.0038

Table 3
Effect of the measurement noise in the steady-state behaviour of the low-power peaking-free high-gain observer (11) with $\ell = 7$. In the table $\tilde{x}_i := \hat{x}_i - x_i$. (a): coloured random noise with band-pass filter. (b): sinusoidal noise with $\omega = 100$. (c): coloured random noise with high-pass filter. (d): sinusoidal noise with $\omega = 5000$.

	(a)	(b)	(b)	(c)
$\|\nu\|_{\mathcal{T}_{0.01}}^{\infty}$	0.298	0.298	4.4894	4.4894
$\|\tilde{x}_1\|_{\mathcal{T}_{0.01}}^{\infty}$	0.078	0.092	0.025	0.028
$\|\tilde{x}_2\|_{\mathcal{T}_{0.01}}^{\infty}$	0.940	1.138	0.317	0.347
$\|\tilde{x}_3\|_{\mathcal{T}_{0.01}}^{\infty}$	5.628	6.919	1.927	2.106
$\|\tilde{x}_4\|_{\mathcal{T}_{0.01}}^{\infty}$	16.687	20.713	5.768	6.307
$\|\tilde{x}_5\|_{\mathcal{T}_{0.01}}^{\infty}$	19.491	24.417	6.797	7.439

Table 4
Effect of the measurement noise in the steady-state behaviour of the high-gain observer (2) with $\ell = 4.17$. In the table $\tilde{x}_i := \hat{x}_i - x_i$. (a): coloured random noise with band-pass filter. (b): sinusoidal noise with $\omega = 100$. (c): coloured random noise with high-pass filter. (d): sinusoidal noise with $\omega = 5000$.

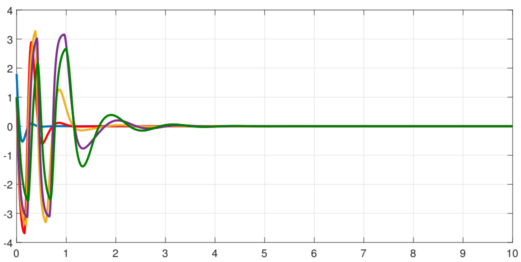

Fig. 3. Behaviour of the error dynamics $\hat{x}_i(t) - x_i(t)$, $i = 1, \ldots, 5$, of system (1), (20) and observer (11).

5 Proofs

5.1 Proof of Theorem 1

The proof follows the same idea of Astolfi and Marconi (2015), with just minor adaptations due to the different dimension of the actual observer ($2n-1$ instead of $2n-2$), and therefore it is just sketched. Let $\tilde{\chi} := \mathrm{col}(\tilde{\chi}_1, \ldots, \tilde{\chi}_n)$ with $\tilde{\chi}_i \in \mathbb{R}^2$ for $i = 1, \ldots, n-1$ and $\tilde{\chi}_n \in \mathbb{R}$, defined as

$$\tilde{\chi}_i := \mathrm{col}\left(\frac{\hat{x}_i - x_i}{\ell^{i-1}}, \frac{\eta_i - x_{i+1}}{\ell^i}\right) \qquad i = 1, \ldots, n-1 \quad (21)$$

and $\tilde{\chi}_n := \ell^{-(n-1)}(\hat{x}_i - x_i)$. By applying the previous change of coordinates to (5), we obtain

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

where the matrix M_n is Hurwitz by design, $\bar{B} = \mathrm{col}(B_{2n-2}, 1)$, $\bar{K}_1 := \mathrm{col}(K_1, 0, \ldots, 0)$, and the function $\Delta_\varphi := \varphi(x, d) - \varphi_s(\hat{x})$ satisfies $|\Delta_\varphi(\tilde{\chi}, x, d)| \le \ell^{n-1}\bar{\varphi}_x|\tilde{\chi}| + R$ for all $(\chi, x, d) \in \mathbb{R}^{2n-1} \times \mathbb{R}^n \times D$, for some $\bar{\varphi}_x > 0$ and $R := 2\max_{\{x \in X, d \in D\}} |\varphi(x, d)|$. By definition, the following bounds $\ell^{-(i-1)}|\hat{x}_i - x_i| \le |\tilde{\chi}|$, $\ell^{-i}|\eta_i - x_{i+1}| \le |\tilde{\chi}|$ and $|\tilde{\chi}| \le |\hat{\boldsymbol{x}} - \boldsymbol{x}|$ hold for $\ell \ge 1$. As a consequence, bounds (7) and (8) can be obtained by using previous bounds and by applying Lemma 4 in Appendix A with $\ell^\star$ indicated in the statement of the theorem. □

5.2 Proof of Proposition 2

Firstly, it is worth expressing the signal (4) as an output of an autonomous system properly initialised. In this respect, we observe that, having defined $S \in \mathbb{R}^{n_w \times n_w}$ and $P \in \mathbb{R}^{1 \times n_w}$ as

$$S := \mathrm{blkdiag}(S_1, \ldots, S_{n_\nu}), \qquad S_i = \begin{pmatrix} 0 & \omega_i \\ -\omega_i & 0 \end{pmatrix},$$

and $P := ((0\ 1)\ (0\ 1)\ \cdots\ (0\ 1))$, the measurement noise (4) can be expressed as output of the following system

$$\begin{aligned} \varepsilon \dot{w} &= S w, \qquad w \in \mathbb{R}^{n_w} \\ \nu &= P w, \end{aligned} \qquad (22)$$

with initial condition $w(0)$ dependent on ν_i^c and ν_i^s. The initial condition $w(0)$, in particular, ranges in a compact set W that is invariant for (22).

Then, the proof of the Proposition conceptually articulates in three parts. In the first part the whole system, given by the observed system (1), the observer (5) and the noise generator (22), is transformed, by means of a coordinates change, into a cascade autonomous system given by an asymptotically stable system (which is the estimation error dynamics) driven by a system with bounded trajectories (which is the parallel of the observed system and of the noise generator). This cascade structure leads to a first conclusion that the state of the estimation error dynamics asymptotically converges to a steady state governed by the state of the driving subsystem (namely x and w). Such a steady state will be clearly affected by the parameter ε characterising (22). The core of the proof is then the characterisation of such a steady state in terms of ε. In this respect, in the second part of the proof an approximation of such a steady state is presented (with an approximation that is of order ε^ρ with ρ properly defined). The result is contained in the forthcoming Lemma 2, which is a technical lemma that can be proved by following the computations in Astolfi, Marconi, Praly and Teel (2016). In the final part of the proof, then, the asymptotic properties of the cascade system are analysed in relation to the approximated steady state to obtain the result claimed in the proposition. The structure of the proof follows the idea originally presented in Astolfi, Marconi, Praly and Teel (2016) for classical high-gain observers (2).

Consider the change of coordinates

$$\tilde{\xi}_i := \mathrm{col}(\hat{x}_i - x_i, \eta_i - x_{i+1}) \qquad i = 1, \ldots, n-1$$

with $\tilde{\xi}_i = \mathrm{col}(\tilde{\xi}_{i1}, \tilde{\xi}_{i2}) \in \mathbb{R}^2$ for all $i = 1, \ldots, n-1$, and $\tilde{\xi}_n := \hat{x}_n - x_n$, that transforms the observer (5) into the form

$$\dot{\tilde{\xi}} \;=\; F\tilde{\xi} + \bar{B}\Delta_\varphi(\tilde{\xi}, x) + G\nu(t) \tag{23}$$

with the matrix F recursively constructed as $F_1 = H_1$,

$$F_i := \begin{pmatrix} F_{i-1} & \bar{N}_i \\ \bar{Y}_i & H_i \end{pmatrix} \quad i = 2, \ldots, n-1, \quad F := \begin{pmatrix} F_{n-1} & 0 \\ \ell\bar{q}_n & -\ell\,\alpha_n \end{pmatrix}$$

with $H_i := A - D_2(\ell)\,K_i\,C$, $Y_i := D_2(\ell)\,K_i\,B^T$, $\bar{N}_i := B_{2(i-1)}B_2^T$, for $i = 1, \ldots, n-1$, $\bar{B} = \mathrm{col}(B_{2n-2}, 1)$, $G := \mathrm{col}\big(G_1, 0, \ldots, 0\big)$ with $G_1 := D_2(\ell)K_1$ and

$$\Delta_\varphi(\tilde{\xi}, x) \;:=\; \varphi_s(\Gamma\tilde{\xi} + x) - \varphi(x, d)\,, \tag{24}$$

with $\Gamma := \mathrm{blkdiag}\,(\underbrace{C_2,\ \ldots,\ C_2}_{(n-1)\text{ times}}, 1)$. 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 (23) and the noise generator (22) read as

$$\begin{aligned} \varepsilon\dot{w} &= Sw \\ \dot{x} &= f(x) \\ \dot{\tilde{\xi}} &= F\tilde{\xi} + B\Delta_\varphi(\tilde{\xi}, x) + GPw\,. \end{aligned} \tag{25}$$

Having taken the parameters $(\underline{\alpha}, \underline{\beta})$ and ℓ 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 Isidori and Byrnes (2008). 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

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

is asymptotically stable for (25). Furthermore, the properties of the high-gain observer when the measurement noise is absent (*i.e.* when $w = 0$) show that $\pi_\varepsilon(0, x) = \{0\}$ for all $x \in X$. 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

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

and let ρ be an arbitrary (integer) number satisfying $\rho \geq m$, with m given by (9). Note that for any n we have $m \geq \upsilon$. The approximation of order ρ of the steady state is then a function $\Psi_\varepsilon : W \times X \to \mathbb{R}^{2n-1}$ defined as

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

in which

$$\begin{aligned} \Psi_i(w, x) &:= \sum_{j=a_i}^{\rho} \psi_{i,j}(w, x)\,\varepsilon^j\,, \quad i = 1, \ldots, n \\ \Lambda_i(w, x) &:= \sum_{j=b_i}^{\rho} \lambda_{i,j}(w, x)\,\varepsilon^j\,, \quad i = 1, \ldots, n-1 \end{aligned} \tag{26}$$

where the $a_i = b_i = i$ for $i = 1, \ldots, \upsilon$, $a_i = n - i + 2$, $b_i := n - i + 1$ for $i = \upsilon + 1, \ldots, n$, with

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

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

Lemma 3 *Consider system* (25) *and the notations introduced before. There exist continuous functions* $\psi_{i,j}(\cdot, \cdot)$ *and* $\lambda_{i,j}(\cdot, \cdot)$ *such that, having defined*

$$\begin{aligned} \mathsf{E}_\varepsilon(w, x) \;:=\; & \frac{\partial\Psi_\varepsilon(w, x)}{\partial w}Sw + \frac{\partial\Psi_\varepsilon(w, x)}{\partial x}f(x) \\ & - F\Psi_\varepsilon(w, x) - GPw - B\Delta_\varphi(\Psi_\varepsilon(w, x), x)\,, \end{aligned}$$

the following holds

$$\begin{aligned} \lim_{\varepsilon \to 0^+} \frac{\mathsf{E}_\varepsilon(w, x)}{\varepsilon^{\rho-1}} = 0 & \qquad \forall\,(w, x) \in W \times X\,, \\ \mathsf{E}_\varepsilon(0, x) = 0 & \qquad \forall\,(\varepsilon, x) \in [0, 1] \times X\,. \end{aligned}$$

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

$$
\begin{aligned}
\psi_{i,a_i}(w,x) &:= \ell^{2i-1}\,\bar{\psi}_{i,a_i}(w,x), & i &= 1,\dots,m,\\
\psi_{i,a_i}(w,x) &:= \ell\,\bar{\psi}_{i,a_i}(w,x), & i &= m+1,\dots,n.
\end{aligned} \tag{27}
$$

Proof: Due to space constraints, we just give here a sketch of the proof, by focusing on the main steps to derive the expression of $\mathsf{E}_\varepsilon(\cdot,\cdot)$. By letting

$$
\mathsf{E}_\varepsilon(\cdot,\cdot) := \mathrm{col}\Big(\mathsf{E}_1,\ \Xi_1,\ \mathsf{E}_2,\ \Xi_2,\ \dots\ ,\ \mathsf{E}_{n-1},\ \Xi_{n-1},\ \mathsf{E}_n\Big)
$$

it can be seen that the E_i, $i = 1,\dots,n$, and Ξ_i, $i = 1,\dots,n-1$, components have the form

$$
\begin{aligned}
\mathsf{E}_1 &= \dot{\Psi}_1 + \ell\alpha_1\Psi_1 - \Lambda_1 - \ell\alpha_1 P w\\
\Xi_1 &= \dot{\Lambda}_1 + \ell^2\beta_1\Psi_1 - \Lambda_2 - \ell^2\beta_1 P w\\
\mathsf{E}_i &= \dot{\Psi}_i + \ell\alpha_i\Psi_i - \Lambda_i - \ell\alpha_i\Lambda_{i-1}\\
\Xi_i &= \dot{\Lambda}_i + \ell^2\beta_i\Psi_i - \Lambda_{i+1} - \ell^2\beta_i\Lambda_{i-1}\\
&\qquad\qquad\qquad\qquad i = 2,\dots,n-2,\\
\mathsf{E}_{n-1} &= \dot{\Psi}_{n-1} + \ell\alpha_{n-1}\Psi_{n-1} - \Lambda_{n-1} - \ell\alpha_{n-1}\Lambda_{n-2}\\
\Xi_{n-1} &= \dot{\Lambda}_{n-1} + \ell^2\beta_{n-1}\Psi_{n-1} - \Delta_\varphi(\overline{\Psi}_\varepsilon, x)\\
&\qquad - \ell^2\beta_{n-1}\Lambda_{n-2}\\
\mathsf{E}_n &= \dot{\Psi}_n + \ell\alpha_n\Psi_n - \Delta_\varphi(\overline{\Psi}_\varepsilon, x) - \ell\alpha_n\Lambda_{n-1}
\end{aligned} \tag{28}
$$

where, for the sake of compactness, we omitted the argument (w,x) from the functions Ψ_i, $i = 1,\dots,n$, Λ_i, $i = 1,\dots,n-1$ and $\overline{\Psi}_\varepsilon := \Gamma\Psi_\varepsilon$. Note that, 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

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

Therefore, we can expand the term Δ_φ by a Taylor series around $\overline{\Psi}_\varepsilon = 0$ to obtain

$$
\Delta_\varphi(\overline{\Psi}_\varepsilon, x) \;=\; \sum_{j=1}^{\rho} \varepsilon^j \phi_j(w,x) + \varepsilon^{\rho+1} R_\varepsilon(w,x).
$$

The main idea of the proof is then 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 ε of order j, with $j = 0,\dots,\rho-1$, for $i = 1,\dots,n$. By considering the term of order 0 in ε in the expression of E_1 and Ξ_1 it is easy to see that

$$
\psi_{1,1}(w,x) = \ell\alpha_1 P S^{-1} w\ , \quad \lambda_{1,1}(w,x) = \ell^2\beta_1 P S^{-1} w\ .
$$

By proceeding iteratively one can select all the functions $\psi_{i,j}$ and $\lambda_{i,j}$ according to the PDEs (28) to show that E_i, Ξ_i are terms in ε^ρ thus satisfying the first part of the lemma. Similarly, the second part of the lemma follows by inspecting the choice of $\psi_{i,j}$ and the PDEs (28). □

With the result of Lemma 3 in hand, we are now in the position of concluding the proof of Proposition 2. Let consider the change of variables

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

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

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

Furthermore, note that

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

and that there exists $\varepsilon_1^\star(\ell) \in (0,1]$ such that [2] for all positive $\varepsilon \le \varepsilon_1^\star(\ell)$

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

As a consequence, we can compute the error dynamics in the new coordinates as

$$
\dot{\zeta} = F\zeta + B\Delta_\varphi(\zeta, \Gamma\Psi_\varepsilon(w,x) + x) + \mathsf{E}_\varepsilon(w,x)\,. \tag{29}
$$

Since the Lipschitz constant of $\Delta_\varphi(\cdot,\cdot)$ is not affected by the value of the arguments, the same values of ℓ that make system (23) ISS with respect to the input $\nu(t)$ make also system (29) ISS with respect to the input $\mathsf{E}_\varepsilon(\cdot,\cdot)$. In particular, there exists $c_0 > 0$ such that

$$
\begin{aligned}
\lim_{t\to\infty}\sup|\zeta(t)| &= \lim_{t\to\infty}\sup|\tilde{\xi}(t) - \Psi_\varepsilon(w(t),x(t))|\\
&\le c_0 \lim_{t\to\infty}\sup|\mathsf{E}_\varepsilon(w(t),x(t))|\\
&\le c_0\,\|\mathsf{E}_\varepsilon(w,x)\|_\infty
\end{aligned}
$$

Using the fact that, for any $\rho \ge m$, $\mathsf{E}_\varepsilon(w,x)$ is a term in ε^ρ, it follows that there exists $c_1 > 0$ such that

$$
\lim_{t\to\infty}\sup|\zeta(t)| \;\le\; c_1\varepsilon^\rho\,\|w\|_\infty\ .
$$

Consider now the expressions of the components $\Psi_i(\cdot,\cdot)$, $i = 1,\dots,n$, of $\Psi_\varepsilon(\cdot,\cdot)$ introduced in Lemma 3. It turns

[2] Note that the value of $\varepsilon^\star$ depends, among other things, on the choice of the set X_δ on which $\varphi_s(\cdot)$ coincides with $\varphi(\cdot)$

out that there exist a positive $\varepsilon_2^\star(\ell) \le \varepsilon_1^\star(\ell)$ and $\mu_2 > 0$ such that

$$\begin{aligned} |\Psi_i(w,x)| &\le c_2\, \varepsilon^i\, \ell^{2i-1}\, |w| \qquad i = 1,\dots,m, \\ |\Psi_i(w,x)| &\le c_2\, \varepsilon^{n-j+2}\, \ell\, |w| \qquad i = m+1,\dots,n, \end{aligned}$$

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

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

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

$$\begin{aligned} \limsup_{t\to\infty} |\hat{x}_i(t) - x_i(t)| &= \limsup_{t\to\infty} |\zeta_{i1}(t) + \Psi_i(w(t),x(t))| \\ &\le c_1 \varepsilon^\rho \|w\|_\infty + c_2 \varepsilon^{n-i+2} \ell \|w\|_\infty \,. \end{aligned}$$

Since $\nu(t) = Pw(t)$, and by recalling that $\|w\|_\infty$ does not depend on the choice of ε, there exists $c_3 > 0$ satisfying $\|w\|_\infty \le c_3 \|\nu\|_\infty$. The result follows by taking an appropriate $\varepsilon^\star(\ell) \le \varepsilon_2^\star(\ell)$ and $\hat{c} > 0$. □

5.3 Proof of Proposition 3

The proof follows the main steps proposed in Astolfi, Marconi and Teel (2016), with just minor adaptations due to the different dimension of the actual observer ($2n-1$ instead of $2n-2$) and the presence of the measurement noise ν, and therefore it is just sketched. Consider the change of coordinates

$$\zeta_i := \mathrm{col}(\hat{x}_i - x_i, \ell^{-1}(\eta_i - x_{i+1})), \qquad i = 1,\dots,n-1$$

and $\zeta_n := \hat{x}_n - x_n$, that transforms system (11) into

$$\begin{aligned} \dot{\zeta}_i &= \ell E_i \zeta_i + \ell^{-1} B_2 u_i + \ell K_i \varpi_i \qquad i = 1,\dots,n-1, \\ \dot{\zeta}_n &= -\ell \alpha_n \zeta_n + u_{n-1} + \ell \alpha_n \varpi_n \end{aligned}$$

where $\varpi_1 = \nu$ and $\varpi_i = \mathrm{sat}_{r_i}(\eta_{i-1})$ for $i = 2,\dots,n$, $u_i := \mathrm{sat}_{r_{i+2}}(\eta_{i+1}) - x_{i+2}$ for $i = 1,\dots,n-2$ and $u_{n-1} := \varphi_s(\hat{x},0) - \varphi(x,d)$. 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 ℓ, such that $\|u_i\|_\infty \le \bar{u}_i$. Furthermore, note that the matrices E_i, $i = 1,\dots,n-1$, are Hurwitz by design of $\underline{\alpha}, \underline{\beta}$ and $\alpha_n > 0$. Recall also that ϖ_i is bounded for any $i = 1,\dots,n$. Hence, by applying Lemma 4 in Appendix A, it turns out that there exist constants $c_{ij} > 0$, with $i = 1,\dots,n$ and $j = 1,\dots,4$, such that

$$|\zeta_i(t)| \le c_{i1} \exp(-c_{i2} \ell\, t) |\zeta_i(0)| + \frac{c_{i3}}{\ell^2} \bar{u}_i + c_{i4}$$

holds for any $\ell \ge 1$ and for $i = 1,\dots,n$. From this, by using the fact that $\frac{1}{\ell}|x_{i+1} - \eta_i| \le |\zeta_i|$ and $|\zeta_i| \le |x_i - \hat{x}_i| + |x_{i+1} - \eta_i|$ hold for all $\ell \ge 1$, the bound (14) immediately follows with $\bar{p}_i := c_{i1}\pi_i + c_{i3}\bar{u}_i + c_{i4}$, with c_{i1} proportional to $\bar{\nu}$ and $\pi_i := \max\limits_{x\in X,\, (\hat{x},\eta)\in \widehat{X}\times E} \{|x_i - \hat{x}_i| + |x_{i+1} - \eta_i|\}$ for $i = 1,\dots,n-1$, and $\pi_n := \max\limits_{x\in X,\, \hat{x}\in \widehat{X}} |x_n - \hat{x}_n|$.

To prove the item (b) of the proposition, we proceed by induction by recursively showing that all the η_i, from $i = 1$ to $i = n-1$, exit from the saturation if the measurement noise is sufficiently small. For this, consider the change of coordinates (21). The dynamics of $\tilde{\chi}_1$ 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 ℓ (for all $\ell \ge 1$). Using the fact that E_1 is Hurwitz, Lemma 5 in Appendix A, 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 $\underline{\ell}_1 \ge 1$ such that for all $\ell \ge \underline{\ell}_1$ and all $\nu(t)$ satisfying $\|\nu\|_\infty < \bar{\nu}_1/\ell$, we have $|\ell\, \tilde{\chi}_1(t)| \le 1$ for all $t \ge T_1$. Hence, by noting that $\ell^{-1}|\eta_1 - x_2| \le |\tilde{\chi}_{12}| \le |\tilde{\chi}_1|$ and $|\eta_1| \le |\eta_1 - x_2| + |x_2|$ holds for any $\ell \ge 1$, we get $|\eta_1(t)| \le \ell|\tilde{\chi}_1(t)| + |x_2(t)| \le r_2 + 1$ for any $\ell \ge \underline{\ell}_1$ and for all $t \ge T_1$, namely $\mathrm{sat}_{r_2}(\eta_1(t)) = \eta_1(t)$ for all $t \ge T_1$. Now we proceed by induction, namely 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 $\ell \ge \underline{\ell}_{i-1}$ and all $\nu(t)$ satisfying $\|\nu\|_\infty < \bar{\nu}_{i-1}/\ell^{i-1}$ then $\mathrm{sat}_{r_{j+1}}(\eta_j(t)) = \eta_j(t)$ for all $j = 1,\dots,i-1$ and $t \ge T_{i-1}$. Let us use the notation $\tilde{\chi}_{[k]} = (\tilde{\chi}_1,\dots,\tilde{\chi}_k)^T$ for the first k-th components of $\tilde{\chi}$. By the Lipschitz mean-value theorem, there exists a $s(\cdot) \in \mathcal{C}_{[0,1]}$ such that $\mathrm{sat}_{r_{i+1}}(\eta_i) - x_{i+2} = s(t)\, \ell^i B_2^T \tilde{\chi}_i$. As a consequence, it turns out that for $t \ge T_{i-1}$ 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 (12), $\bar{K}_1 := (K_1, 0, \dots, 0)^T$ and $u_i := \mathrm{sat}_{r_{i+2}}(\eta_{i+1}) - x_{i+2}$ for $i = 1,\dots,n-2$. Note that the initial condition $\varepsilon_{[i]}(0)$ ranges in a compact set $\mathcal{E}_i$ not dependent on ℓ (for all $\ell \ge 1$). By Lemma 5 in Appendix A applied with $k = i$ and $X = \mathcal{E}_i$, it turns out that there exist a $\bar{\nu}_i > 0$ and, for all $T_i > T_{i-1}$, a $\underline{\ell}_i \ge \underline{\ell}_{i-1}$ such that, for all $\ell \ge \underline{\ell}_i$ and all $\nu(t)$ satisfying $\|\nu\|_\infty < \bar{\nu}_i/\ell^i$, the inequality $|\ell^i \tilde{\chi}_{[i]}(t)| \le 1$ holds for all $t \ge T_i$. From this, by noting that $\ell^{-i}|\eta_i - x_{i+1}| \le |\tilde{\chi}_{i2}| \le |\tilde{\chi}_{[i]}|$ holds for any $\ell \ge 1$ and and $|\eta_i| \le |\ell^i \tilde{\chi}_{[i]}| + |x_{i+1}|$ it follows that for all $\ell \ge \underline{\ell}_i$ and any $\nu(t)$ fulfilling $\|\nu\|_\infty < \bar{\nu}_i/\ell^i$, the inequality $|\eta_i(t)| \le \ell^i |\tilde{\chi}_{[i]}| + |x_{i+1}| \le r_{i+1} + 1$ holds for all

$t \geq T_i$, namely $\text{sat}_{r_{i+1}}(\eta_i(t)) = \eta_i(t)$ for all $t \geq T_i$. This completes the proof. □

6 Conclusion

We presented a new class of nonlinear high-gain observers. 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 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, see e.g., Astolfi et al. (2017). 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. The effect of high-frequency measurement noise has been analysed in the nonlinear context with the approach introduced in Astolfi, Marconi, Praly and Teel (2016). We modelled the measurement noise as a summation of sinusoidal signals and we analysed the steady-state behaviour of the observer by computing an approximation of a partial differential equation. Numerical simulations confirm the validity of the approach. The proposed analysis tool is of its own interest since it can be applied to other classes of nonlinear observers to analyse the effect of measurement noise.

References

Ahrens, J. H. and Khalil, H. K. (2009). High-gain observers in the presence of measurement noise: A switched-gain approach, Automatica **45**(4): 936–943.

Astolfi, D. (2016). Observers and robust output regulation for nonlinear systems, PhD thesis, PSL Research University.

Astolfi, D., Isidori, A. and Marconi, L. (2017). Output regulation via low-power construction, Feedback Stabilization of Controlled Dynamical Systems, Springer, pp. 143–165.

Astolfi, D. and Marconi, L. (2015). A high-gain nonlinear observer with limited gain power, IEEE Trans. Autom. Control **60**(11): 3059–3064.

Astolfi, D., Marconi, L., Praly, L. and Teel, A. R. (2016). Sensitivity to high-frequency measurement noise of nonlinear high-gain observers, IFAC-PapersOnLine **49**(18): 862–866.

Astolfi, D., Marconi, L. and Teel, A. R. (2016). Low-power peaking-free high-gain observers for nonlinear systems, Eur. Control Conf., pp. 1424–1429.

Astolfi, D. and Praly, L. (2017). Integral action in output feedback for multi-input multi-output nonlinear systems, IEEE Trans. Autom. Control **62**(4): 1559–1574.

Atassi, A. N. and Khalil, H. K. (2000). Separation results for the stabilization of nonlinear systems using different high-gain observer designs, Syst. Control Lett. **39**(3): 183–191.

Boizot, N., Busvelle, E. and Gauthier, J.-P. (2010). An adaptive high-gain observer for nonlinear systems, Automatica **46**(9): 1483–1488.

Byrnes, C. I. and Isidori, A. (2004). Nonlinear internal models for output regulation, IEEE Transactions on Automatic Control **49**(12): 2244–2247.

Gauthier, J.-P. and Kupka, I. (2001). Deterministic observation theory and applications, Cambr. Univ. Press.

Isidori, A. and Byrnes, C. I. (2008). Steady-state behaviors in nonlinear systems with an application to robust disturbance rejection, Annual Reviews in Control **32**(1): 1–16.

Khalil, H. K. (2017). Cascade high-gain observers in output feedback control, Automatica **80**: 110–118.

Khalil, H. K. and Praly, L. (2014). High-gain observers in nonlinear feedback control, Int. J. Robust Nonlin. Control **24**(6): 993–1015.

Khalil, H. K. and Priess, S. (2016). Analysis of the use of low-pass filters with high-gain observers, IFAC-PapersOnLine **49**(18): 488–492.

Lancaster, P. and Rodman, L. (1995). Algebraic riccati equations, Clarendon press.

Maggiore, M. and Passino, K. M. (2003). A separation principle for a class of non-uco systems, IEEE Trans. Autom. Control **48**(7): 1122–1133.

Martinez-Guerra, R. and Mata-Machuca, J. L. (2013). Fault detection and diagnosis in nonlinear systems: a differential and algebraic viewpoint, Springer.

Sanfelice, R. G. and Praly, L. (2011). On the performance of high-gain observers with gain adaptation under measurement noise, Automatica **47**(10): 2165–2176.

Sprott, J. C. (2010). Elegant chaos: algebraically simple chaotic flows, World Scientific.

Teel, A. and Praly, L. (1994). Global stabilizability and observability imply semi-global stabilizability by output feedback, Systems & Control Letters **22**: 313–325.

Teel, A. R. (2016). Further variants of the Astolfi/Marconi high-gain observer, Am. Control Conf., pp. 993–998.

Wang, L., Astolfi, D., Marconi, L. and Su, H. (2017). High-gain observers with limited gain power for systems with observability canonical form, Automatica **75**: 16–23.

A Appendix

The proofs of the forthcoming lemmas follow by Lyapunov arguments that are omitted.

Lemma 4 *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* ν, *and with* k *and* ℓ *positive numbers. Suppose that:*

(i) there exists $P = P^T > 0$ *such that* $PA(s) + A(s)^T P \leq -I$ *holds for all* $s \in [0,1]$;

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

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

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

Lemma 5 *Consider the system*

$$\dot{x} \;=\; \ell A(s(t))x + \frac{1}{\ell^k}Bu(t) + \ell\, K\,\nu(t) \tag{A.1}$$

where $s \in \mathcal{C}_{[0,1]}$, *with state* $x \in \mathbb{R}^n$, *and with* k *and* ℓ *positive numbers. Assume there exists a* $\bar{u} > 0$ *such that* $\|u\|_\infty < \bar{u}$ *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* $\ell > \underline{\ell}$ *and any* $s(\cdot) \in \mathcal{C}_{[0,1]}$, *trajectories of* (A.1) *originating from* X *and subject to disturbances* $\nu(t)$ *fulfilling* $\|\nu\|_\infty \leq \bar{\nu}/\ell^k$, *satisfy* $|\ell^k\, x(t)| \leq 1$ *for all* $t \geq T$.