On the Performance of High-gain Observers with Gain Adaptation under Measurement Noise ⋆

Ricardo G. Sanfelice [a] Laurent Praly [b]

[a] *Department of Aerospace and Mechanical Engineering, University of Arizona 1130 N. Mountain Ave, AZ 85721, USA (Tel: 520-626-0676; e-mail: sricardo@u.arizona.edu).*

[b] *CAS, ParisTech, Ecole des Mines, 35 rue Saint Honoré, 77305, Fontainebleau, France (e-mail: Laurent.Praly@ensmp.fr).*

Abstract

We address the problem of state observation for a system whose dynamics may involve poorly known, perhaps even nonlocally Lipschitz functions and whose output measurement may be corrupted by noise. It is known that one way to cope with all these uncertainties and noise is to use a high-gain observer with a gain adapted on-line. The proposed method, while presented for a particular case, relies on a "generic" analysis tool based on the study of differential inequalities involving quadratic functions of the error system in two coordinate frames plus the gain adaptation law. We establish that, for bounded system solutions, the estimated state and the gain are bounded. Moreover, we provide an upper bound for the mean value of the error signals as a function of the observer parameters. Since due to perturbations the gain adaptation law may drive the observer/plant interconnection to nearby the boundary of its stability region, oscillatory behavior may emerge. To overcome this issue, we suggest an adaptive procedure based on a space averaging technique involving several copies of the observer.

1 Introduction

We consider nonlinear systems in the form [1]

$$
\begin{aligned}
\dot{z} &= f_z(x_1,\ldots,x_n,z,t),\\
\dot{x}_1 &= x_2 + f_1(x_1,z,t),\\
\dot{x}_2 &= x_3 + f_2(x_1,x_2,z,t),\\
&\vdots\\
\dot{x}_{n-1} &= x_n + f_{n-1}(x_1,x_2,\ldots,x_{n-1},z,t),\\
\dot{x}_n &= f_n(x_1,x_2,\ldots,x_n,z,t),\\
y &= x_1 + m.
\end{aligned}
\tag{1}
$$

For such systems, we are interested in estimating the components x_1 to x_n of any solution that is bounded in positive times.

To that end, we propose a high-gain observer with adaptive gain that measures the plant's output y perturbed by m and is given by

$$
\begin{aligned}
\dot{\hat{x}}_1 &= \hat{x}_2 + \hat{f}_1(\hat{x}_1,t) - k_1 r(\hat{y}-y),\\
\dot{\hat{x}}_2 &= \hat{x}_3 + \hat{f}_2(\hat{x}_1,\hat{x}_2,t) - k_2 r^2(\hat{y}-y),\\
&\vdots\\
\dot{\hat{x}}_{n-1} &= \hat{x}_n + \hat{f}_{n-1}(\hat{x}_1,\ldots,\hat{x}_{n-1},t) - k_{n-1}r^{n-1}(\hat{y}-y),\\
\dot{\hat{x}}_n &= \hat{f}_n(\hat{x}_1,\hat{x}_2,\ldots,\hat{x}_n,t) - k_n r^n(\hat{y}-y),\\
\hat{y} &= \hat{x}_1,\\
\dot{r} &= \phi(r, y-\hat{y}),
\end{aligned}
$$

where the functions $\hat{f}_i$ and the positive constants k_i, which are the nominal gains, are to be chosen, r is the observer's gain, which is introduced to increase the nominal gain if needed, and ϕ defines the adaptation law.

The domain of application of traditional, constant high-gain observers ([11,10]) has been enlarged by incorporating dynamic gain adaptation; see, e.g., [15,6,16,3,2]. Dynamic gain adaptation is reminiscent of what has been proposed in the adaptive control literature for online tuning of control parameters; see, e.g., [9,14,12,18]. When it is known that the gain r should be larger than some function of the state that is observable (see [3,2,21,20] for instance), then it is easy to design a sat-

[1] The time dependence allows the presence of inputs.

isfactory gain adaptation law. When we only know the effect or the properties that r can guarantee when it is large enough, (see [15,6,16,3]-[7]), then it is more difficult to design an adaptation law guaranteeing robust performance. Indeed, typically this adaptation is such that the gain r is nondecreasing along solutions. Unfortunately, it is known in various contexts that such a gain adaptation may lead to serious growth problems when perturbations such as measurement noise are present (see, e.g., [9, Example 4.2], [19, Figure 6.a], and [18]). A wide variety of fixes have been proposed in the literature to stop r from increasing without bound. For instance, there exist the dead-zone [9,19] or λ-tracking approach [17], the sigma modification [13], and, more recently, in the context of output feedback stabilization, the hybrid approach proposed in [23] consisting of decreasing (increasing) r by resetting it to a smaller (larger) value when the output of the system decreases (respectively, increases). The point is that, instead of keeping the gain r at large values when it is not needed, more sophisticated mechanisms that tune r to the local (in time) plant's data are needed in real-world applications. In fact, it has been established in [25,4] that for the constant high-gain case, measurement noise introduces an upper limit for the gain when good performance is taken into account. Gain adaptation laws aiming at satisfying this objective have been proposed recently in [1,5]. In [1], the authors are in a context in which a bound on the f_i's in (1) is known. This allows them to let the gain r switch between two appropriate values depending on the magnitude of the error $\hat{y} - y$. In [5], the authors design an adaptation law for r relying on the knowledge of an upper bound for r and of the Lipschitz constant of the f_i's.

We design an adaptation law ϕ for r that does not require information on the upper bound for r nor of the Lipschitz constant of the f_i's. Our approach consists of analyzing the following set of inequalities resulting from the interconnection between the plant in (1) and the observer proposed above:

$$\begin{aligned} \frac{\dot{V}_r(\varepsilon)}{r} &\leq -\alpha_1(r)V_r(\varepsilon) + \beta_1(r), \\ \dot{r} &= \phi(r, y - \hat{y}), \\ \frac{\dot{V}_s(\xi)}{s} &\leq -\theta_1 V_s(\xi) + \theta_2 + \alpha_2(r)(y - \hat{y})^2, \\ \beta_2(r)\,(x_1 - \hat{x}_1)^2 \;\leq\; V_r(\varepsilon) &\leq \alpha_3(s)V_s(\xi). \end{aligned} \tag{2}$$

The functions V_r and V_s are quadratic in ε and ξ, respectively, while $\dot{V}_r$ and $\dot{V}_s$ are their derivatives along solutions, where ε and ξ are two different coordinates obtained from the same error state $e := \hat{x} - x$. The functions α_1, α_2, and α_3 are increasing whereas β_1 and β_2 are decreasing. The constants θ_1 and θ_2 are positive, and s is a positive analysis parameter. Particular constructions of these functions are given in Section 4.3. With these definitions, (2) induces the following mechanism. From the last inequality, if V_r is large, then V_s is also large. This is possible only if $\alpha_2(r)(y - \hat{y})^2$ has been large for some time as the third inequality indicates. If it was r that was large, then, with the first inequality, using the monotonicity properties of α_1 and β_1, this contradicts that V_r is large. So it has to be that $\|\hat{y} - y\|$ is large. If ϕ takes positive values when $\|\hat{y} - y\|$ is large, then, from the second inequality, r will also become large, forcing V_r to decrease via the first inequality. Since this does not put any constraint on ϕ when $\|\hat{y} - y\|$ is small, our idea is to let ϕ take nonpositive values in such case.

The paper is organized as follows. Section 2 presents the construction of the observer as well as the main result. Its proof along that of technical lemmas are in Section 4. Section 3 is devoted to the presentation of an illustrating academic example.

Notation

For notation convenience, we utilize the following symbols and definitions throughout the paper:

- $\widetilde{K} := [k_1 \;\, k_2 \;\, \ldots \;\, k_n]^\top$, where $k_i \in \mathbb{R}$ for all $i \in \{1, 2, \ldots, n\}$.
- $\mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn})$ denotes the diagonal matrix with entries a_{ii}, $i = 1, 2, \ldots, n$.
- $\Lambda(r) = \mathrm{diag}(r, \ldots, r^n)$.
- $N_{n-1} = \mathrm{diag}(0, 1, \ldots, n-1)$.
- Given $b \in \mathbb{R}$, define $R = bI + N_{n-1}$.
- $\widetilde{R}(r, s) = \mathrm{diag}\left(1 - \left(\frac{r}{s}\right), 1 - \left(\frac{r}{s}\right)^2, \ldots, 1 - \left(\frac{r}{s}\right)^n\right)$.
- Given $x \in \mathbb{R}^n$, $\|x\|$ denotes the Euclidean norm of x.
- Given $A \in \mathbb{R}^{n \times n}$, $\|A\|$ denotes the induced 2-norm of A.
- Given a function $t \mapsto f(t)$, $\|f\|_\infty$ denotes $\mathrm{esssup}_t \|f(t)\|$.
- Given a matrix $P \in \mathbb{R}^{n \times n}$, $\lambda_{\min}(P)$ and $\lambda_{\max}(P)$ denote the minimum and maximum values of its eigenvalues, respectively.

2 Observer expression and main result

System (1) can be compactly written as

$$\begin{aligned} \dot{z} &= f_z(x_1, \ldots, x_n, z, t), \\ \dot{x} &= Ax + F(x, z, t), \\ y &= x_1 + m, \end{aligned} \tag{3}$$

where

$$A := \begin{bmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 1 \\ 0 & 0 & 0 & \dots & \dots & 0 \end{bmatrix},$$

$$F(x,z,t) := \begin{bmatrix} f_1(x_1,z,t) \\ f_2(x_1,x_2,z,t) \\ \vdots \\ f_{n-1}(x_1,x_2,\dots,x_{n-1},z,t) \\ f_n(x_1,x_2,\dots,x_n,z,t) \end{bmatrix},$$

$(z,x) \in \mathbb{R}^m \times \mathbb{R}^n$ is the plant's state, $y \in \mathbb{R}$ is the perturbed plant's output, and m represents the noise in the measurements of x_1.

We study the high-gain observer discussed in Section 1 for (1) with the particular gain adaptation law defined by

$$\phi(r, \hat{y} - y) := p_1 \left(((\hat{y} - y)^2 - p_2) r^{1-2b} + \frac{p_2}{r^{2n}} \right),$$

with p_1 and p_2 parameters to be chosen positive and b to be taken in $(0, \frac{1}{2})$. As discussed in Section 1, it is such that the gain r increases at least when $(\hat{y} - y)^2$ is larger than p_2, but it decreases when $(\hat{y} - y)^2$ is smaller than $p_2 \left(1 - \frac{1}{r^{2n+1-2b}}\right)$. Note that our adaptation law makes the interval $[1, +\infty)$ forward invariant for the r-component of any solution.

The above expression for ϕ has some resemblance with the one corresponding to an update law with dead zone; cf. [9,19]. More precisely, in the most standard case and in our context, an update law with dead zone would assume the form

$$\dot{r} = p_1 \max\left\{0, (\hat{y} - y)^2 - p_2\right\} r^{1-2b}, \tag{4}$$

in which case, $\dot{r}$ is always nonnegative.

With the definitions above, the proposed observer for the components x_1 to x_n of (1) becomes

$$\dot{\hat{x}} = A\hat{x} + \hat{F}(\hat{x}, t) - K(r)(\hat{y} - y), \tag{5}$$

$$\dot{r} = p_1 \left(((\hat{y} - y)^2 - p_2) r^{1-2b} + \frac{p_2}{r^{2n}} \right), \tag{6}$$

$$\hat{y} = \hat{x}_1, \tag{7}$$

where $\hat{x} \in \mathbb{R}^n$, $\hat{y} \in \mathbb{R}$,

$$\hat{F}(\hat{x},t) := \begin{bmatrix} \hat{f}_1(\hat{x}_1,t) \\ \hat{f}_2(\hat{x}_1,\hat{x}_2,t) \\ \vdots \\ \hat{f}_{n-1}(\hat{x}_1,\hat{x}_2,\dots,\hat{x}_{n-1},t) \\ \hat{f}_n(\hat{x}_1,\hat{x}_2,\dots,\hat{x}_n,t) \end{bmatrix},$$

and with the notation $K(r) := \Lambda(r)\widetilde{K}$. Given $b \in (0, \frac{1}{2})$ and using [21, Lemma 1], a vector $\widetilde{K} \in \mathbb{R}^n$ can be chosen to guarantee the existence of real numbers d_0 and d_1, and of a symmetric matrix P such that

$$0 < d_0, \quad 0 < d_1, \quad 0 < P, \tag{8}$$

$$\left(A - \widetilde{K}C\right)^\top P + P\left(A - \widetilde{K}C\right) \le -2d_0 P, \tag{9}$$

$$\frac{b}{2}P \le RP + PR \le d_1 P, \tag{10}$$

where $C := [1\ 0\ 0\ \dots\ 0] \in \mathbb{R}^n$. Unless otherwise stated, the parameter b of the gain adaptation law is constrained to the set $(0, \frac{1}{2})$.

To establish our main result, we require F and $\hat{F}$ to satisfy the following property:

> **Property (∗):** For each compact set $\mathfrak{C} \subset \mathbb{R}^m \times \mathbb{R}^n$, there exist $\gamma, L \in \mathbb{R}^n$ satisfying, for each $i \in \{1, 2, \dots, n\}$ and all (x, w, z) such that $(z, x + w) \in \mathfrak{C}$,
>
> $$\begin{aligned} \| f_i(x_1 + w_1, x_2 + w_2, \dots, x_i + w_i, z, t) - \\ \hat{f}_i(x_1, x_2, \dots, x_i, t) \| \le \gamma_i + L_i \textstyle\sum_{j=1}^{i} \|w_j\|. \end{aligned} \tag{11}$$
>
> for almost all t.

In particular, the constant vector γ captures a bound on the unmodeled dynamics, both in the dynamics defined by the functions f and $\hat{f}$, while L corresponds to a bound on the Lipschitz constant of the mismatch between these functions.

The following lemma introduces conditions for which this property is guaranteed.

Lemma 2.1 *Assume the function F is such that the function $(x, z) \mapsto F(x, z, t)$ is locally bounded uniformly in t and the function $\hat{F}$ is bounded. Under this condition, Property (∗) holds.*

Proof: For a compact set $\mathfrak{C}$, local boundedness of $(x, z) \mapsto f_i(x, z, t)$ uniformly in t implies the existence of $\gamma_i' > 0$ such that $\|f_i(x_1 + e_1, \dots, x_i + e_i, z, t)\| \le \gamma_i'$ for all

$(z, x+e) \in \mathfrak{C}$ and for all t. Boundedness of $\hat{f}_i$ implies the existence of $\gamma_i'' > 0$ such that $\|\hat{f}_i(x_1, \ldots, x_i, t)\| \leq \gamma_i''$ for all $(x_1, \ldots, x_i)$ and for all t. Then, the claim follows with $\gamma_i \geq \gamma_i' + \gamma_i''$, $L_i \geq 0$, $i \in \{1, 2, \ldots, n\}$. Note that this proof indicates that one could pick $L_i = 0$ in (11). But keeping L_i gives more flexibility and less conservative results when a Lipschitz property holds. ■

Next, we state our main result. For any pair $(F, \hat{F})$ such that Property $(*)$ holds, it establishes that, for each complete [2] and bounded solution to the plant (3) and bounded measurement noise, the interconnection between the plant and the proposed observer, which results in the system (3),(5)-(6), is such that no finite escape time occurs and that solutions are bounded. Moreover, it provides an explicit bound for the mean value of the error signals.

Theorem 2.2 *Assume the pair $(F, \hat{F})$ is such that Property $(*)$ holds. Assume further that $\hat{F}$ satisfies the Carathéodory conditions [3]. Given $b \in (0, \frac{1}{2})$, let $\widetilde{K}, d_0, d_1$ and P satisfy (9). Then, for each real number $M_\infty \geq 0$ and positive gain adaptation law parameters satisfying*

$$p_1 > 0, \quad p_2 \geq 4M_\infty^2 \left(1 + \frac{2\,\lambda_{\max}(P) \left\|\widetilde{K}\right\|^2}{{d_0}^2 \lambda_{\min}(P)}\right) \tag{12}$$

we have that, for each

A) Carathéodory solution $t \mapsto (z(t), x(t))$ to (3) that is complete and bounded,
B) Measurement noise given by a measurable function $t \mapsto m(t)$ satisfying $\|m\|_\infty \leq M_\infty$, and
C) Initial condition $(\hat{x}(0), r(0))$ of (5)-(6) with $r(0) \geq 1$,

the corresponding Carathéodory solutions $t \mapsto (z(t), x(t), \hat{x}(t), r(t))$ to system (3),(5)-(6)

(1) Exist and are complete,
(2) Are bounded on $[0, +\infty)$, and
(3) Satisfy

$$\limsup_{T \to +\infty} \frac{1}{T} \int_t^{t+T} (\hat{y}(\tau) - y(\tau))^2 d\tau \;\leq\; p_2 \quad \forall t \geq 0, \tag{13}$$

$$\limsup_{T \to +\infty} \limsup_{t \to +\infty} \frac{1}{T} \int_t^{t+T} |\hat{x}_i(\tau) - x_i(\tau)|^2 d\tau \leq B_{i,\circ}(p_1, p_2) \tag{14}$$

for all $i \in \{1, 2, \ldots, n\}$, where $B_{i,\circ}(p_1, p_2) > 0$ is given in (15); see Remark 2.4.

Remark 2.3 *Property $(*)$ provides an upper bound on the mismatch $F - \hat{F}$ for all $(x, \hat{x}, z, r, t)$ on compact sets for the (z, x) components. Further measurability and continuity conditions on $\hat{F}$ and m guarantee local existence of Carathéodory solutions to system (5)-(6), once a solution of (3) is given. Note that the assumptions do not guarantee that complete and bounded Carathéodory solutions $t \mapsto (z(t), x(t))$ to (3) exist. In fact, such solutions can fail to exist, even locally. Theorem 2.2 asserts properties only for solutions $t \mapsto (z(t), x(t), \hat{x}(t), r(t))$ to system (3),(5)-(6) associated to a complete and bounded Carathéodory solution $t \mapsto (z(t), x(t))$ to (3).*

Remark 2.4 *While expression (13) suggests that the bound for the mean value of the output error can be made small by picking p_2 small, the bound in (14) requires that p_2 satisfies (12). That is, the bound in (14) is constrained by the size of the measurement noise and the conditions (8)-(10). The bounds in (13)-(14) provide an estimate of achieved performance, in which $B_{i,\circ}(p_1, p_2)$ is given by*

$$\begin{aligned} B_{i,\circ}(p_1, p_2) \;=\; \min_{s > s^*} \Bigg\{ & \frac{s^{2i}\left(a_1(s) + 2c_1\, M_\infty^2\right)}{\lambda_{\min}(P)\left(d_0 s^2 - c_0\, \hat{L}\right)} \\ & + \frac{2c_1 \left(B_1(s, p_1, p_2) + \dfrac{1}{p_2} B_2\right)^{\frac{2n}{1-2b}}}{s^{2(n-i)} \lambda_{\min}(P)\left(d_0 s^2 - c_0\, \hat{L}\right)} p_2 \Bigg\}, \end{aligned} \tag{15}$$

where $c_0 := \frac{2\,\lambda_{\max}(P)}{d_0}$, $c_1 := c_0 \left\|\widetilde{K}\right\|^2$, $\hat{L} := \frac{2}{\lambda_{\min}(P)} \sum_{i=1}^n i\, L_i^2$, $a_1(s) := 2\, c_0 \sum_{i=1}^n \frac{\gamma_i^2}{s^{2i}}$, $a_2(s) := \max\left\{s^b, s^{(b+n-1)}\right\}^2 \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}$, $s^* := \max\left\{\frac{p_1 p_2 (2n+1)}{d_0} + \sqrt{\frac{c_0 \hat{L}}{d_0}}, \bar{s}, 1\right\}$,

$$B_1(s, p_1, p_2) := \widetilde{B}_1(s, p_1, p_2)^{1-2b} + \left(2d_1 + \frac{2c_0\, \hat{L}}{p_1 p_2}\right) \widetilde{B}_1(s, p_1, p_2) + 2,$$

[2] A solution is complete if its domain of definition is $[0, +\infty)$

[3] A function $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ satisfies the Carathéodory conditions if: $x \mapsto f(x,t)$ is continuous uniformly in t; $t \mapsto f(x,t)$ is measurable uniformly in x; for each compact set $A = \mathcal{X} \times [a,b] \subset \mathbb{R}^n \times \mathbb{R}$, there exists a function $m_A : [a,b] \to \mathbb{R}_{\geq 0}$, Lebesgue integrable on $[a,b]$, such that $\|f(x,t)\| \leq m_A(t)$ for all $(x,t) \in A$.

$$\widetilde{B}_1(s,p_1,p_2) := \max\left\{ p_1\left(\frac{4a_2(s)c_1M_\infty^2 + a_2(s)\,a_1(s)}{s^{2b}d_0\lambda_{\min}(P)\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right)}\right) - p_1p_2\left(\frac{2a_2(s)c_1}{s^{2b}d_0\lambda_{\min}(P)\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right)}\right) + \left(\frac{4c_1\,s^{1-2b}\,a_2(s)}{d_0\lambda_{\min}(P)(2n+1)}\right)\left(2\frac{\left(d_1p_1p_2 + c_0\,\hat{L}\right)}{d_0}\right)^{\frac{2n+1}{1+2b}}, \; 4p_1\frac{a_1(1)+c_1M_\infty^2}{d_0^2\lambda_{\min}(P)}\right\},$$

$B_2 := \frac{4}{d_0\lambda_{\min}(P)}a_1(1)$, where, for fixed parameters p_2 and γ, the constant $\overline{s} > 0$ is constrained to satisfy $2c_1p_2 > a_1(\overline{s})$, which is always possible by picking $\overline{s}$ large enough. Note that $B_{i,\circ}(p_1,p_2)$ is given by the minimization of the sum of two terms. The first term is the bound that one would obtain if the constant vector L were known and the gain r were kept constant, and satisfying $r > \max\left\{\sqrt{\frac{c_0\,\hat{L}}{d_0}}, 1\right\}$. Indeed, in this case, only the first term of (15) remains, that is,

$$\limsup_{T\to+\infty}\limsup_{t\to+\infty}\frac{1}{T}\int_t^{t+T}|\hat{x}_i(\tau)-x_i(\tau)|^2d\tau \le \frac{r^{2i}(a_1(r)+2c_1M_\infty^2)}{\lambda_{\min}(P)\left(d_0r^2 - c_0\,\hat{L}\right)}$$

for all $i\in\{1,2,\ldots,n\}$. The second term in $B_{i,\circ}$ corresponds to the effect of the gain adaptation law. Moreover, when the bound on the mismatch $F-\hat{F}$ in Property (*) is such that γ is zero, since a_1 and B_2 are identically zero in such case, using the bound on p_2 given in (12), the bound $B_{i,\circ}$ can be rewritten as

$$B_{i,\circ}(p_1,p_2) = \min_{s\,>\,s^*}\frac{1}{2}\frac{s^{2i}c_1}{\lambda_{\min}(P)\left(d_0s^2 - c_0\,\hat{L}\right)}\left(\frac{1}{\dfrac{c_1}{d_0\lambda_{\min}(P)}+1} + \frac{4\,B_1(s,p_1,p_2)^{\frac{2n}{1-2b}}}{s^{2n}}\right)p_2. \tag{16}$$

The following corollary of Theorem 2.2 follows from Remark 2.4 when γ is zero.

Corollary 2.5 *Under the assumptions of Theorem 2.2, given $b\in(0,\frac{1}{2})$ and letting $\widetilde{K}, d_0, d_1$ and P satisfy (9), if $\gamma = 0$ then, for each real number $M_\infty \ge 0$, parameters of the gain adaptation law (6) satisfying (12), there exist a constant $\beta > 0$ such that for each Carathéodory solution $t\mapsto(z(t),x(t))$ to (3), measurement noise m, and initial condition $(\hat{x}(0),r(0))$ satisfying conditions A), B), and C) of Theorem 2.2, respectively, the corresponding Carathéodory solutions $t\mapsto(z(t),x(t),\hat{x}(t),r(t))$ to system (3),(5)-(6) satisfy*

$$\limsup_{T\to+\infty}\limsup_{t\to+\infty}\frac{1}{T}\int_t^{t+T}\|\hat{x}(\tau)-x(\tau)\|^2d\tau \le \beta\,M_\infty^2. \tag{17}$$

Remark 2.6 *Corollary 2.5 follows from the fact that when $\gamma = 0$ we have a_1 and B_2 identically zero. In such a case, we obtain from (16)*

$$\beta := \frac{2c_1}{\lambda_{\min}(P)}\min_{s\,>\,s^*}\frac{\left(\max\limits_{i\in\{1,2,\ldots,n\}} s^{2i}\right)}{\left(d_0s^2 - c_0\,\hat{L}\right)}\left(1 + 4\left(1+\frac{c_1}{d_0\lambda_{\min}(P)}\right)\frac{B_1(s,p_1,p_2)^{\frac{2n}{1-2b}}}{s^{2n}}\right)$$

when $p_2 = 4M_\infty^2\left(1+\frac{c_1}{d_0\lambda_{\min}(P)}\right)$.

Furthermore, in the absence of measurement noise, the next corollary of Theorem 2.2 follows immediately from the expression of the bound (14). In fact, when γ and m are zero, the first term in (15) vanishes and the bound can be made arbitrarily small by picking p_2 small.

Corollary 2.7 *Under the assumptions of Theorem 2.2, given $b\in(0,\frac{1}{2})$ and letting $\widetilde{K}, d_0, d_1$ and P satisfy (9), if $\gamma = 0$ and $m\equiv 0$ then, for every $\widetilde{\varepsilon} > 0$, there exists $\bar{p}_2 > 0$ such that, for each parameters of the gain adaptation law (6) satisfying $p_1 > 0, p_2\in(0,\bar{p}_2]$, each Carathéodory solution $t\mapsto(z(t),x(t))$ to (3), and each initial condition $(\hat{x}(0),r(0))$ satisfying conditions A) and C) of Theorem 2.2, respectively, the corresponding Carathéodory solutions $t\mapsto(z(t),x(t),\hat{x}(t),r(t))$ to system (3),(5)-(6) satisfy*

$$\limsup_{T\to+\infty}\limsup_{t\to+\infty}\frac{1}{T}\int_t^{t+T}\|\hat{x}(\tau)-x(\tau)\|^2d\tau \le \widetilde{\varepsilon}. \tag{18}$$

The proofs of the above results are in Section 4.

3 A Numerical Example

To illustrate the main features of our dynamic high-gain observer it is already sufficient to consider an elementary second order linear system. Consider the linear plant

$$\dot{x}_1 = x_2 + \nu_1\,x_1 + \nu_2,\quad \dot{x}_2 = 0,\quad y = x_1 + m, \tag{19}$$

with $\nu_1,\nu_2 > 0$, ν_1 being known, but ν_2 unknown. Note that the plant can be rewritten as in (3) with $F(x) =$

$[\nu_1\, x_1 + \nu_2\ \ 0]^\top$ and that $x = [-\frac{\nu_2}{\nu_1}\ \ \ 0]^\top$ is an equilibrium. Following Section 2, the observer (5) is designed with $\hat{F}(\hat{x}) = [\nu_1\, \hat{x}_1\ \ 0]^\top$ and is given by

$$
\begin{aligned}
\dot{\hat{x}}_1 &= \hat{x}_2 + \nu_1\, \hat{x}_1 - k_1\, r\, (\hat{y} - y), \quad \dot{\hat{x}}_2 \;=\; -k_2\, r^2 (\hat{y} - y), \\
\dot{r} &= p_1 \left(((\hat{y} - y)^2 - p_2) r^{1-2b} + \frac{p_2}{r^4} \right), \quad \hat{y} = \hat{x}_1 .
\end{aligned}
\tag{20}
$$

With this particular choice, it follows that Property $(*)$ holds with $\gamma = [\nu_2\ \ 0]^\top$ and $L = [\nu_1\ \ 0]^\top$. Straightforward calculations show that (8)-(10) holds, in particular, for the following set of parameters: $d_0 = 0.37$, $d_1 = 2.66$, $b = 0.183$, $P = \begin{bmatrix} 0.75 & -0.50 \\ -0.50 & 0.88 \end{bmatrix}$, $\widetilde{K} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $p_1 = 20$. In the analysis to follow, we study the effect of varying the parameter ν_1 and the measurement noise m for different values of p_2. The parameter ν_2 is fixed at 0.1.

First, we consider the case where $\nu_1 = 1.9$ and m is an independently and identically normally distributed stochastic process with mean 0.02 and standard deviation 0.015. The histogram of this noise is shown in Figure 1, where the dashed vertical lines correspond to $\pm\sqrt{p_2}$ whereas the dotted vertical line corresponds to the x_1 component $-\frac{\nu_2}{\nu_1}$ of the solution. It follows that the probability of the norm of the noise m to be larger than $\sqrt{p_2}$ is small but non zero.

A simulation of (19)-(20) with $p_2 = 0.0025$ and initial conditions $x(0) = [-\frac{\nu_2}{\nu_1}\ \ 0]^\top$, $\hat{x}(0) = [1\ \ 0]^\top$, and $r(0) = 1$ is shown in Figure 2. In dark/blue, it shows the first 300 sec of the trajectory of the gain r and the tail of the $\hat{x}_2 - x_2$ component of the resulting simulation, while in light gray/magenta, it shows the simulation with the dead-zone law in (4). A numerical comparison of the dead-zone and the proposed gain adaptation laws is given in the Table 1 ($\overline{\hat{x}_1 - x_1}$ denotes the mean value). As expected, the proposed gain adaptation law

method	$\overline{\hat{x}_1 - x_1}$	std $\hat{x}_1 - x_1$	$\overline{\hat{x}_2 - x_2}$	std $\hat{x}_2 - x_2$
dead-zone	-1.99e-02	3.98e-03	1.38e-01	1.38e-02
proposed	-1.99e-02	6.58e-03	1.38e-01	9.00e-03

Table 1
Numerical comparison of the dead-zone and the proposed gain adaptation laws with $p_2 = 0.0025$ and $\nu_1 = 1.9$.

yields a gain $r(t)$ that decreases while guaranteeing the estimates to converge, but the dead-zone law uses a gain with asymptotic value that is nearly six times larger. This is due to a large error $\hat{y} - y$ during the transient stage or a potentially bad choice of the initial condition $r(0)$ (for which there is no a priori information on how to select it). As pointed out in [25,4], keeping the gain at large values may compromise performance when measurement noise is present. This is what the numbers in Table 1 indicate.

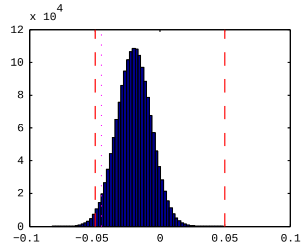
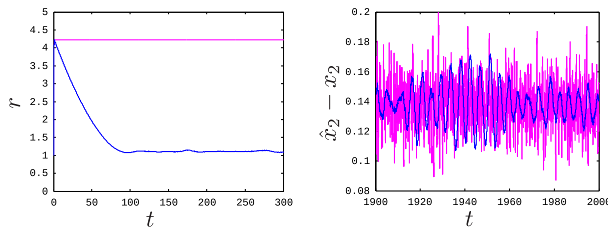


Fig. 1. Measurement noise histogram. The dashed vertical lines correspond to $\pm\sqrt{p_2}$ while the dotted vertical line corresponds to the x_1 component $-\frac{\nu_2}{\nu_1}$ of the solution.



(a) r versus time (first 300sec). (b) $\hat{x}_2 - x_2$ versus time (tail).

Fig. 2. Complete trajectory of r and final part of $\hat{x}_2 - x_2$ for the system (19) and our observer (20) (dark/blue) and for the one with dead-zone in (4) (light gray/magenta) when $p_2 = 0.0025$, $\nu_1 = 1.9$ and $m = \mathcal{N}(0.02, 0.015)$.

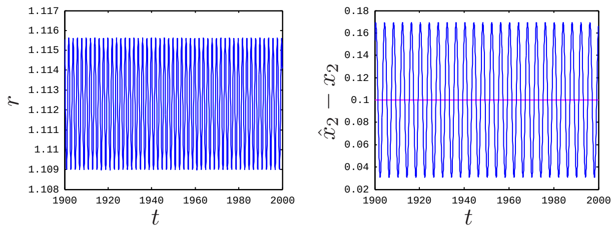
The analysis sketched in Section 1 to argue about boundedness does not rule out the possibility of oscillations in $\hat{x}$ and r. In fact, up to now, a goal of our adaptation law was to bring the gain r back to one. But an unitary value for r corresponds to an unstable equilibrium point of the error system. In fact, note that when there is no measurement noise, the error system is given by

$$
\begin{aligned}
\dot{e}_1 &= e_2 + \nu_1\, e_1 - k_1\, r\, e_1 - \nu_2, \\
\dot{e}_2 &= -k_2\, r^2\, e_1, \\
\dot{r} &= p_1 \left((e_1^2 - p_2) r^{1-2b} + \frac{p_2}{r^4} \right).
\end{aligned}
\tag{21}
$$

The point $e_1 = 0$, $e_2 = \nu_2$, $r = 1$ is an equilibrium to this system. Around this point, the linearization matrix is

$$
A := \left. \begin{pmatrix} \nu_1 - k_1 r & 1 & 0 \\ -k_2 r^2 & 0 & 0 \\ 0 & 0 & -p_1 p_2 (5 - 2b) \end{pmatrix} \right|_{r=1}, \tag{22}
$$

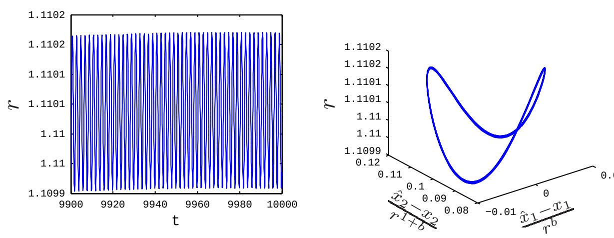
which is Hurwitz if and only if $\nu_1 < k_1$. This condition is satisfied in the simulation described above and depicted in Figure 2. But, when $\nu_1 \geq k_1$, unsatisfactory behavior may appear. Indeed, for $\nu_1 = 2.22$, since $k_1 = 2$, instead of obtaining the trajectories shown in Figure 2, our gain adaptation law leads to the oscillatory behavior shown in Figure 3, where only the tail of both $\hat{x}_2 - x_2$ and r are shown. This is confirmed by the numbers in Table 2.



(a) r versus time (tail). (b) $\hat{x}_2 - x_2$ versus time (tail).

Fig. 3. Trajectories of r and $\hat{x}_2 - x_2$ (tails) when $p_2 = 0.0025$, $\nu_1 = 2.22$ and $m = 0$.

Along the same lines of Corollary 2.7, the size of the oscillations can be reduced by appropriately tuning the observer parameters. For instance, for $p_2 = 0.0001$, which corresponds to the original value of p_2 divided by 25, the numerical results are given in Table 3. The tail of r as well as the trajectories in the $\left(\frac{\hat{x}_1 - x_1}{r^b}, \frac{\hat{x}_2 - x_2}{r^{1+b}}, r\right)$-space for the case where $p_2 = 0.0001$, $\nu_1 = 2.22$ and $m = 0$ are show in Figure 4. As expected, we have a (compact)



(a) r versus time (tail). (b) Set $\mathcal{A}$ in the $\left(\frac{\hat{x}_1 - x_1}{r^b}, \frac{\hat{x}_2 - x_2}{r^{1+b}}, r\right)$-space.

Fig. 4. Trajectory of r (tail) and the asymptotic phase portrait of $\left(\frac{\hat{x}_1 - x_1}{r^b}, \frac{\hat{x}_2 - x_2}{r^{1+b}}, r\right)$ when when $p_2 = 0.0001$, $\nu_1 = 2.22$ and $m = 0$.

Ω-limit set in 3-D space. But, very importantly, r is oscillating around the value $\frac{\nu_1}{k_1} = 1.11$, which would make the matrix A in (22) marginally stable. A rather simple solution to the oscillatory problem consists of adapting online the appropriate value we want r to converge to. We call this value the *nominal gain* and denote it $\overline{r}$. Unfortunately, this value is likely to evolve with time as the

method	$\overline{\hat{x}_1 - x_1}$	std $\hat{x}_1 - x_1$	$\overline{\hat{x}_2 - x_2}$	std $\hat{x}_2 - x_2$
dead-zone	1.62e-11	1.54e-09	1.00e-01	1.08e-08
proposed	-1.22e-04	3.12e-02	1.00e-01	4.90e-02

Table 2
Numerical comparison of the dead-zone and the proposed gain adaptation laws with $p_2 = 0.0025$ and $\nu_1 = 2.22$.

method	$\overline{\hat{x}_1 - x_1}$	std $\hat{x}_1 - x_1$	$\overline{\hat{x}_2 - x_2}$	std $\hat{x}_2 - x_2$
dead-zone	3.57e-11	1.31e-08	1.00e-01	1.17e-07
proposed	1.05e-05	6.20e-03	1.00e-01	9.72e-03

Table 3
Numerical comparison of the dead-zone and the proposed gain adaptation laws with $p_2 = 0.0001$ and $\nu_1 = 2.22$.

system solution evolves, making an adaptive procedure based only on the time evolution of r not appropriate. For instance, a time average could be misleading when the oscillations are created by the system solution itself.

An alternative view of the situation is to consider that, in the oscillatory case, the observer answer is, at each time, not a single point, but rather the set of points $\mathcal{A}$ depicted in Figure 4(b). The knowledge at each time of what this set is and the mean value for its r component is could provide a good estimation of the limiting value for the nominal value $\overline{r}$. A way to learn the set $\mathcal{A}$ is to sample it, i.e., to have a sufficient (but finite) number of points moving on this set that are as far apart as possible in such a way that their distribution represents the set $\mathcal{A}$ itself well enough. Then, the proposed approach is to have several copies of the observer with estimates that are sufficiently far apart from each other running simultaneously as to provide the desired points moving on the (compact) set $\mathcal{A}$. Then, the objective is to solve an optimization problem involving the chordal distance as the cost function, e.g.,

$$\max_{\hat{x}_{i,\cdot} \in \mathcal{A}} \min_{i \neq j} |\hat{x}_{i,\cdot} - \hat{x}_{j,\cdot}|,$$

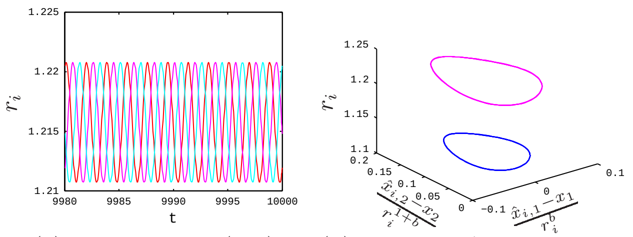
where the $\hat{x}_{i,\cdot}$'s are the state estimates given by copies of the observer. This problem is closely related to *packing*[4] (see [8,24] for instance). To solve this problem, we propose to inject a (small) disturbance in the observer dynamics in the direction of the gradient of the above cost. With a possible theoretical analysis in mind and in view of the technicalities presented in the next section, we propose the following collection of n_{obs} observers:

$$\begin{aligned}
\dot{\hat{x}}_{i,1} &= \hat{x}_{i,2} + \nu_1\, \hat{x}_{i,1} - k_1\, r_i\, (\hat{y}_i - y) \\
&\quad + \overline{k}_1 r_i^{1+b} \sum_{j \neq i} \left(\frac{\hat{y}_i - y}{r_i^b} - \frac{\hat{y}_j - y}{r_j^b} \right), \\
\dot{\hat{x}}_{i,2} &= -k_2\, r_i^2\, (\hat{y}_i - y) + \overline{k}_2 r_i^{2+b} \sum_{j \neq i} \left(\frac{\hat{y}_i - y}{r_i^b} - \frac{\hat{y}_j - y}{r_j^b} \right), \\
\dot{r}_i &= p_1 \left(((\hat{y}_i - y)^2 - p_2) r_i^{1-2b} + \frac{p_2}{r_i^4} \right), \quad \hat{y}_i = \hat{x}_{i,1},
\end{aligned} \tag{23}$$

$i = 1, \ldots, n_{obs}$, where the rightmost terms in the first two equations correspond to the injection terms with gains $\overline{k}_1$ and $\overline{k}_2$ obtained as $\begin{pmatrix} \overline{k}_1 \\ \overline{k}_2 \end{pmatrix} = \mu P^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, with μ to be chosen large enough to speed up the sampling of $\mathcal{A}$ while keeping the observer stable. With $n_{obs} = 3$, $\mu = 0.026$, $p_2 = 0.0025$, $\nu_1 = 2.22$ and $m = 0$, the resulting set $\mathcal{A}$ with the proposed collection of observers is depicted in light gray/magenta in Figure 5(b) while the

[4] But we are facing the extra problem of not knowing the "manifold" where the points evolve.

corresponding r_i's are shown in Figure 5(a). It also shows in dark/blue the set $\mathcal{A}$ for the initial observer, which indicates that the effect of the repellent terms in the collection of observers is to increase the mean value of the r_i's. Note that, now, the r_i's are oscillating around 1.21, which would make the linearization matrix associated with the error system given in (22) marginally stable, and that their phase difference is approximately $\frac{2\pi}{3}$ rad (meaning that the separation of the points sampling $\mathcal{A}$ is indeed maximized). From the collection of observers



(a) r versus time (tail). (b) Set $\mathcal{A}$ in the $\left(\frac{\hat{x}_1-x_1}{r^b}, \frac{\hat{x}_2-x_2}{r^{1+b}}, r\right)$-space.

Fig. 5. Trajectories of the r_i's (tails) and set $\mathcal{A}$ in the $\left(\frac{\hat{x}_{i,1}-x_1}{r_i^b}, \frac{\hat{x}_{i,2}-x_2}{r_i^{1+b}}, r_i\right)$-space when $n_{obs} = 3$, $\mu = 0.026$, $p_2 = 0.0025$, $\nu_1 = 2.22$ and $m = 0$.

we extract, at each time t, the average value

$$\overline{r}(t) = \frac{1}{3}\sum_{i=1}^{3} r_i(t), \tag{24}$$

which we consider to be the right nominal value for the gain of our initial observer. Our motivation for using this nominal value for r emerges from the following conjecture we draw from our analysis: the appropriate value of $\overline{r}$ is in the convex hull of the values that r takes from $\mathcal{A}$. On the other hand, averaging the estimates $\hat{x}_{i,\cdot}$ to get a better estimate may not be a good idea since there is no guarantee that the $x_{i,\cdot}$'s would be in the convex hull of the estimates.

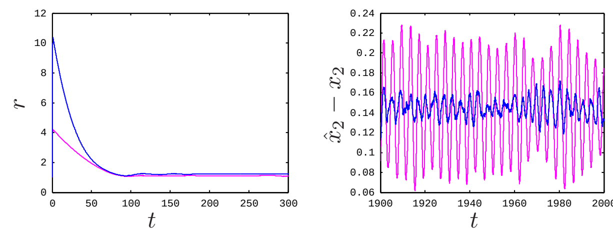
Our initial observer in (5)-(7) can be rewritten in terms of the nominal gain $\overline{r}$ by replacing the original r by the product $r\,\overline{r}$ and noting that the properties established earlier for the original observer still hold when $\overline{r}$ is constant. Then, we obtain

$$\begin{aligned} \dot{\hat{x}}_1 &= \hat{x}_2 + \nu_1 \hat{x}_1 - k_1\, r\, \overline{r}\, (\hat{y}-y), \\ \dot{\hat{x}}_2 &= -k_2\, r^2\, \overline{r}^2\, (\hat{y}-y), \\ \dot{r} &= p_1\left(((\hat{y}-y)^2 - p_2) r^{1-2b} + \frac{p_2}{r^4}\right), \quad \hat{y} \;=\; \hat{x}_1. \end{aligned} \tag{25}$$

Then, for the same setting as in the simulations presented in Figure 3 but with measurement noise, i.e., with $n_{obs} = 3$, $\mu = 0.026$, $p_2 = 0.0025$, $\nu_1 = 2.22$, and m an independently and identically normally distributed stochastic process with mean 0.02 and standard deviation 0.015, the trajectories for r and $\hat{x}_2 - x_2$ obtained with $\overline{r}$ as in (24) are shown in Figure 6 (dark/blue). The simulation for our original observer (25) is also shown (light gray/magenta). We observe that, for the observer in (25), the oscillations are reduced significantly, as is confirmed by the numbers in Table 4.

method	$\overline{\hat{x}_1 - x_1}$	std $\hat{x}_1 - x_1$	$\overline{\hat{x}_2 - x_2}$	std $\hat{x}_2 - x_2$
dead-zone	-2.03e-02	4.30e-03	1.45e-01	1.57e-02
proposed	-2.03e-02	2.93e-02	1.45e-01	4.59e-02
modified	-2.02e-02	5.40e-03	1.45e-01	8.61e-03

Table 4
Numerical comparison of the dead-zone, proposed, and modified gain adaptation laws with $p_2 = 0.0025$ and $\nu_1 = 2.22$.



(a) r versus time (tail). (b) $\hat{x}_2 - x_2$ versus time (tail).

Fig. 6. Trajectories of r and $\hat{x}_2 - x_2$ (tails) when $p_2 = 0.0025$, $\nu_1 = 2.22$ and $m = 0$.

4 Proofs

Theorem 2.2 is about system (3), (5)-(6) with state $\mathfrak{x} = (z, x, \hat{x}, r)$, whose dynamics are compactly written in the form

$$\dot{\mathfrak{x}} = \mathfrak{f}(\mathfrak{x}, t). \tag{26}$$

For any C^1 function $\mathfrak{x} \mapsto h(\mathfrak{x})$, its Lie derivative $\mathcal{L}_{\mathfrak{f}} h$ is $\mathcal{L}_{\mathfrak{f}} h = \frac{\partial h}{\partial \mathfrak{x}}(\mathfrak{x})\,\mathfrak{f}(\mathfrak{x}, t)$. It is useful to distinguish this derivative with the time derivative denoted by "$\cdot$". In particular for a Carathéodory solution $t \mapsto \mathfrak{x}(t)$ to (26), we have $\dot{\overbrace{h(\mathfrak{x}(t))}} = \mathcal{L}_{\mathfrak{f}} h(\mathfrak{x}(t))$, but, in general, only for almost all t in the domain of definition of the solution.

4.1 Error dynamics

With the error state vector definition $e = \hat{x} - x$, we obtain

$$\mathcal{L}_{\mathfrak{f}} e = Ae + \hat{F}(\hat{x}, t) - F(x, z, t) - K(r)e_1 + K(r)m\ . \tag{27}$$

By following [20] for instance, we introduce the following r-scaled error coordinates

$$\varepsilon := \frac{1}{r^{b-1}} \Lambda(r)^{-1} e \tag{28}$$

or equivalently, for each $i \in \{1, 2, \ldots, n\}$, $\varepsilon_i = \frac{e_i}{r^{b+i-1}}$. We obtain

$$\begin{aligned}\mathcal{L}_{\mathfrak{f}}\varepsilon_i &= \frac{\mathcal{L}_{\mathfrak{f}}e_i}{r^{b+i-1}} - (b+i-1)\varepsilon_i \frac{\mathcal{L}_{\mathfrak{f}}r}{r} \\ &= r\varepsilon_{i+1} - k_i r\varepsilon_1 - \frac{\mathcal{L}_{\mathfrak{f}}r}{r}(b+i-1)\varepsilon_i \\ &+ \frac{\hat{f}_i(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_i, t) - f_i(x_1, x_2, \ldots, x_i, z, t)}{r^{b+i-1}} + \frac{k_i}{r^{b-1}}m.\end{aligned}$$

In compact form, this reads

$$\frac{\mathcal{L}_{\mathfrak{f}}\varepsilon}{r} = \left(A - \widetilde{K}C - R\frac{\mathcal{L}_{\mathfrak{f}}r}{r^2}\right)\varepsilon + \Delta(x, \hat{x}, z, r, t) + \frac{\widetilde{K}}{r^b}m, \tag{29}$$

where $C = [1 \ 0 \ \ldots \ 0]^\top \in \mathbb{R}^n$ and Δ is the function defined as

$$\Delta(x, \hat{x}, z, r, t) := \frac{1}{r^b}\Lambda(r)^{-1}(\hat{F}(\hat{x}, t) - F(x, z, t)). \tag{30}$$

The following bound on Δ will be used to derive the right-hand sides of the first and third inequality in (2). For a proof, see [22].

Lemma 4.1 *Assume Property (∗) holds. Then, for each compact set $\mathfrak{C} \subset \mathbb{R}^m \times \mathbb{R}^n$, and for all $(x, \hat{x}, z, r, t)$ such that $(z, x) \in \mathfrak{C}$, we have*

$$\|\Delta(x, \hat{x}, z, r, t)\|^2 \leq \frac{2}{r^{2b}}\sum_{i=1}^{n}\frac{\gamma_i^2}{r^{2i}} + \frac{\hat{L}}{r^2}\varepsilon^\top P\varepsilon, \tag{31}$$

where $\hat{L}$ is given by $\hat{L} = \frac{2}{\lambda_{\min}(P)}\sum_{i=1}^{n} i\, L_i^2$.

Along with the r-scaled error coordinates ε, we introduce s-scaled error coordinates, where s is a positive constant parameter (only for analysis, not part of the observer) which remains to be chosen. More precisely, let

$$\xi := \frac{1}{s^{b-1}}\Lambda(s)^{-1}e. \tag{32}$$

Then, its Lie derivative satisfies

$$\begin{aligned}\mathcal{L}_{\mathfrak{f}}\xi_i &= s\xi_{i+1} - \frac{k_i}{s^{b-1}}\left(\frac{r}{s}\right)^i(\hat{y} - y) \\ &+ \frac{\hat{f}_i(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_i, t) - f_i(x_1, x_2, \ldots, x_i, z, t)}{s^{b+i-1}},\end{aligned}$$

where, by definition, $\hat{y} - y = e_1 - m$. This can be written compactly as

$$\frac{\mathcal{L}_{\mathfrak{f}}\xi}{s} = \left(A - \widetilde{K}C\right)\xi + \Delta(x, \hat{x}, z, s, t) + \frac{\widetilde{K}}{s^b}m + \frac{\widetilde{R}\widetilde{K}}{s^b}(\hat{y} - y),$$

where we recall the notation

$$\widetilde{R}(r, s) = \operatorname{diag}\left(1 - \left(\frac{r}{s}\right), 1 - \left(\frac{r}{s}\right)^2, \ldots, 1 - \left(\frac{r}{s}\right)^n\right).$$

Note that we have the following inequality, which we shall use later on:

$$\|\widetilde{R}(r, s)\| \leq \left\|1 - \left(\frac{r}{s}\right)^n\right\| \qquad \forall s, r > 0. \tag{33}$$

Also, by proceeding as in the proof of Lemma 4.1, we have for all $s \geq 1$ and $(x, \hat{x}, z, s, t)$ such that $(z, x) \in \mathfrak{C}$,

$$\|\Delta(x, \hat{x}, z, s, t)\|^2 \leq \frac{2}{s^{2b}}\sum_{i=1}^{n}\frac{\gamma_i^2}{s^{2i}} + \frac{\hat{L}}{s^2}\xi^\top P\xi. \tag{34}$$

4.2 Properties of Quadratic Functions of e

With $P = P^\top > 0$ satisfying (9) and (10), we define

$$V_r(\varepsilon) := \varepsilon^\top P\varepsilon \quad , \qquad V_s(\xi) := \xi^\top P\xi \ .$$

From the definitions of ε and ξ, we get

$$V_r(\varepsilon) \leq a_2(s)V_s(\xi) \quad \forall e \in \mathbb{R}^n,\ r \geq 1,\ b \in (0, 1/2),\ s > 0, \tag{35}$$

where $a_2(s) = \max\left\{s^b, s^{(b+n-1)}\right\}^2 \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}$. We have

$$\lambda_{\min}(P)\,\varepsilon_1^2 \leq V_r(\varepsilon) \qquad\qquad \forall \varepsilon \in \mathbb{R}^n. \tag{36}$$

The following properties of V_r and V_s are key in establishing our main result.

Lemma 4.2 *Assume Property (∗) holds. Then, for each compact set $\mathfrak{C} \subset \mathbb{R}^m \times \mathbb{R}^n$, $V_r(\varepsilon)$ satisfies, for all $(x, \hat{x}, z, r)$ such that $(z, x) \in \mathfrak{C}$ and almost all t, the following property:*

$$\begin{aligned}\frac{\mathcal{L}_{\mathfrak{f}}V_r}{r} \leq &-\left(d_0 - \left(\frac{d_1\, p_1 p_2}{r^{1+2b}} + \frac{c_0\,\hat{L}}{r^2}\right)\right)V_r(\varepsilon) \\ &+ \frac{a_1(r)}{r^{2b}} + \frac{c_1}{r^{2b}}m^2,\end{aligned} \tag{37}$$

where the constants γ and L are obtained, for the given compact set $\mathfrak{C}$, from Lemma 4.1.

Proof: We obtain

$$\begin{aligned}\frac{\mathcal{L}_{\mathfrak{f}}V_r}{r} &= \varepsilon^\top\left(P\left(A - \widetilde{K}C\right) + \left(A - \widetilde{K}C\right)^\top P\right)\varepsilon \\ &+ 2\varepsilon^\top P\frac{\widetilde{K}}{r^b}m - \varepsilon^\top\left(PR + R^\top P\right)\varepsilon\frac{\mathcal{L}_{\mathfrak{f}}r}{r^2} + 2\varepsilon^\top P\Delta(\cdot).\end{aligned}$$

By decomposing $\mathcal{L}_{\mathfrak{f}} r = r_{dot+} + r_{dot-}$, where $r_{dot+} = \max\{\mathcal{L}_{\mathfrak{f}} r, 0\}$ and $r_{dot-} = \min\{\mathcal{L}_{\mathfrak{f}} r, 0\}$, using (9) and Lemma 4.1, and completing squares, we get

$$\begin{aligned}\frac{\mathcal{L}_{\mathfrak{f}} V_r}{r} \leq & -2d_0 \varepsilon^\top P\varepsilon - \frac{b}{2}\frac{r_{dot+}}{r^2}\varepsilon^\top P\varepsilon - d_1 \frac{r_{dot-}}{r^2}\varepsilon^\top P\varepsilon \\ & + \frac{d_0}{2}\varepsilon^\top P\varepsilon + \frac{2\lambda_{\max}(P)}{d_0}\|\Delta(x,\hat{x},z,r,t)\|^2 + \frac{d_0}{2}\varepsilon^\top P\varepsilon \\ & + \frac{2\lambda_{\max}(P)}{d_0}\left\|\frac{\widetilde{K}}{r^b} m\right\|^2 .\end{aligned}$$

Combining terms, using (31) and (6) yields (37). ■

Lemma 4.3 *Assume Property (∗) holds. For each compact set $\mathfrak{C} \subset \mathbb{R}^m \times \mathbb{R}^n$, $V_s(\xi)$ satisfies, for all $(x,\hat{x},z,r)$ such that $(z,x) \in \mathfrak{C}$ and almost all t, the following property:*

$$\begin{aligned}\frac{\mathcal{L}_{\mathfrak{f}} V_s}{s} \leq & -\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right) V_s(\xi) + \frac{a_1(s)}{s^{2b}} + 2\frac{c_1}{s^{2b}} m^2 + \\ & 2\frac{c_1}{s^{2b}}\left(1 - \left(\frac{r}{s}\right)^n\right)^2 (\hat{y}-y)^2,\end{aligned} \tag{38}$$

where the constants γ and L are obtained, for the given compact set $\mathfrak{C}$, from Lemma 4.1.

Proof: We have

$$\begin{aligned}\frac{\mathcal{L}_{\mathfrak{f}} V_s}{s} = & \xi^\top \left(P\left(A - \widetilde{K}C\right) + \left(A - \widetilde{K}C\right)^\top P\right)\xi \\ & + 2\xi^\top P\Delta(x,\hat{x},z,s,t) + 2\xi^\top P\frac{\widetilde{K}}{s^b} m \\ & + 2\xi^\top P \frac{\widetilde{R}(r,s)\widetilde{K}}{s^b}(e_1 - m).\end{aligned}$$

Proceeding as in the proof of Lemma 4.2, and using (34), it follows

$$\begin{aligned}\frac{\mathcal{L}_{\mathfrak{f}} V_s}{s} \leq & -\left(d_0 - \frac{c_0 \hat{L}}{s^2}\right)\xi^\top P\xi + \frac{2c_0}{s^{2b}}\sum_{i=1}^n \frac{\gamma_i^2}{s^{2i}} \\ & + \frac{2c_0}{s^{2b}}\left\|\widetilde{K}\right\|^2 m^2 + \frac{2c_0}{s^{2b}}\left\|\widetilde{R}(r,s)\widetilde{K}\right\|^2 (e_1 - m)^2 .\end{aligned}$$

The claim follows using the bound (33). ■

4.3 Proof of Theorem 2.2

Since $\hat{F}$ satisfies the Carathéodory conditions by assumption, we are guaranteed that, to any Carathéodory solution to (3) which is defined and bounded on $[0,+\infty)$, each measurement noise satisfying $\|m\|_\infty \leq M_\infty$, and each initial condition $(\hat{x}(0), r(0))$, with $r(0) \geq 1$, there corresponds a (maybe nonunique) Carathéodory solution $t \mapsto (z(t), x(t), \hat{x}(t), r(t))$ to (3),(5)-(6) defined on some right maximal interval $[0,\sigma)$. Our task here is to prove that σ is infinite, i.e., the solution is complete, and that the solution is bounded on $[0,+\infty)$ and satisfies (13)-(18).

For each Carathéodory solution $t \mapsto (z(t), x(t))$ to (3) that is defined and bounded on $[0,+\infty)$ there exists a compact set $\mathfrak{C}$ such that $(z(t), x(t)) \in \mathfrak{C}$ for all $t \geq 0$. This is the (solution dependent) compact set to be used in our technical results in Section 4.1 and Section 4.2. It follows that we have $\mathfrak{C}$-dependent functions $\hat{L}$ and a_1 such that, by combining (37), (6), (38), (35), and (36), we have, for almost all t in $[0,\sigma)$,

$$\begin{aligned}\frac{\dot{V}_r(\varepsilon(t))}{r(t)} \leq & -\left(d_0 - \left(\frac{d_1 p_1 p_2}{r(t)^{1+2b}} + \frac{c_0\,\hat{L}}{r(t)^2}\right)\right) V_r(\varepsilon(t)) \\ & + \frac{a_1(r(t))}{r(t)^{2b}} + \frac{c_1}{r(t)^{2b}} m(t)^2, \\ \dot{r}(t) = & \, p_1\left(((\hat{y}(t) - y(t))^2 - p_2) r(t)^{1-2b} + \frac{p_2}{r(t)^{2n}}\right), \\ \frac{\dot{V}_s(\xi(t))}{s} \leq & -\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right) V_s(\xi(t)) + \frac{a_1(s)}{s^{2b}} \\ & + 2\frac{c_1}{s^{2b}} m(t)^2 + 2\frac{c_1}{s^{2b}}\left(1 - \left(\frac{r(t)}{s}\right)^n\right)^2 (\hat{y}(t) - y(t))^2,\end{aligned}$$

and, for all t in $[0,\sigma)$,

$$\frac{\lambda_{\min}(P)}{r^{2b}}(x_1(t) - \hat{x}_1(t))^2 \;\leq\; V_r(\varepsilon(t)) \;\leq\; a_2(s) V_s(\xi(t)). \tag{39}$$

Using the constant bound on m, these inequalities correspond to a particular case of those in (2). They capture the main feature of the dynamic interconnection between the plant (1) and the proposed observer. We have sketched in the introduction how they can be exploited to prove Theorem 2.2. We proceed in four steps, which are presented in the next sections.

4.3.1 No finite escape time in ε and r

Let $k \geq 0$ to be fixed later. Using (37) and (6), we have, for almost all t in $[0,\sigma)$,

$$\begin{aligned}\overbrace{kV_r(\varepsilon(t)) + r(t)}^{\cdot} \leq & -k\left(d_0\, r(t) - \left(\frac{d_1 p_1 p_2}{r(t)^{2b}} + \frac{c_0\,\hat{L}}{r(t)}\right)\right) V_r \\ & + k\, r(t)^{1-2b} a_1(r(t)) + k\, c_1\, r(t)^{1-2b} m(t)^2 \\ & + \left(p_1(\hat{y}(t) - y(t))^2 - p_1 p_2\right) r(t)^{1-2b} + \frac{p_1 p_2}{r(t)^{2n}}\end{aligned}$$

Then, with (36),

$$p_1(\hat{y}(t)-y(t))^2 r(t)^{1-2b} \leq \frac{2p_1 r(t)}{\lambda_{\min}(P)} V_r(\varepsilon(t)) \;+\; 2p_1 m(t)^2 r(t)^{1-2b} \;.$$

Since the L^∞ norm of m is bounded by M_∞, the function $r \mapsto r^{1-2b} a_1(r)$ is nonincreasing, and $r(t) \geq 1$ for all t in $[0,\sigma)$, by picking $k \geq \frac{2p_1}{d_0\,\lambda_{\min}(P)}$, we obtain for almost all t in $[0,\sigma)$

$$\begin{aligned}\overbrace{kV_r(\varepsilon(t))+r(t)}^{\cdot} \leq\;& k\left(d_1p_1p_2+c_0\,\hat{L}\right)V_r(\varepsilon(t))+p_1p_2\\ &+r(t)^{1-2b}\left(k\,c_1\,\|m\|_\infty^2+2\|m\|_\infty^2 p_1-p_1p_2\right)+ka_1(1)\\ \leq\;& \max\left\{\left(d_1p_1p_2+c_0\,\hat{L}\right),(k\,c_1+2\,p_1)\,\|m\|_\infty^2-p_1p_2\right\}\times\\ &\qquad (kV_r(\varepsilon(t))+r(t))+ka_1(1)+p_1p_2.\end{aligned} \tag{40}$$

This establishes that $V_r(\varepsilon(t))$ and $r(t)$ cannot grow faster than exponentially. Since the solution $t \mapsto (x(t),z(t))$ is known to be defined on $[0,+\infty)$, with the definition of ε, we conclude by contradiction that the solution $t \mapsto (x(t),z(t),\hat{x}(t),r(t))$ of the system (3),(5)-(6) is also complete, i.e., σ is infinite.

4.3.2 Boundedness of ε

To prove boundedness of ε, we select s large enough. Since for $s \geq 1$, we have

$$\left\|1-\left(\frac{r}{s}\right)^n\right\| \;\leq\; r^n \;, \tag{41}$$

by multiplying both sides of (6) by $\frac{c_1}{p_1\;s^{2b}} r^{2n}$ and solving for the term with factor $(\hat{y}-y)^2$, we get that

$$\begin{aligned}2\frac{c_1}{s^{2b}}\left\|1-\left(\frac{r}{s}\right)^n\right\|^2(\hat{y}-y)^2 &\leq 2\frac{c_1}{s^{2b}}r^{2n}r^{1-2b}(\hat{y}-y)^2\\ &= 2\frac{c_1}{p_1s^{2b}}r^{2n}\mathcal{L}_{\mathfrak{f}}r+2\frac{c_1p_2}{s^{2b}}r^{2n}r^{1-2b}-2\frac{c_1p_2}{s^{2b}}\;.\end{aligned}$$

With (38), this leads to, for almost all $t \geq 0$,

$$\begin{aligned}&\overbrace{\frac{V_s(\xi(t))}{s}-\frac{c_1}{p_1s^{2b}}\frac{2}{2n+1}r(t)^{2n+1}}^{\cdot} \leq \\ &-s\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)\left(\frac{V_s(\xi(t))}{s}-\frac{c_1}{p_1s^{2b}}\frac{2}{2n+1}r(t)^{2n+1}\right)+\\ &\frac{a_1(s)}{s^{2b}}+\frac{2c_1p_2}{s^{2b}}r(t)^{2n}r^{1-2b}-\frac{2c_1p_2}{s^{2b}}+\frac{2c_1}{s^{2b}}m(t)^2-\\ &s\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)\frac{c_1}{p_1s^{2b}}\frac{2}{2n+1}r(t)^{2n+1}.\end{aligned} \tag{42}$$

Now observe that $\frac{p_1p_2(2n+1)}{d_0}+\sqrt{\frac{c_0\hat{L}}{d_0}}$ is the largest solution of

$$s^2d_0-s\,p_1p_2(2n+1)-c_0\,\hat{L}=0\;. \tag{43}$$

It follows that picking $s > \max\left\{\frac{p_1p_2(2n+1)}{d_0}+\sqrt{\frac{c_0\hat{L}}{d_0}}\;,\,1\right\}$ implies that

$$\frac{c_1p_2}{s^{2b}}r^{2n+1}\left(\frac{1}{r^{2b}}-\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)\frac{s}{p_1p_2(2n+1)}\right)\leq 0.$$

Then, inequality (42) becomes, for almost all $t \geq 0$,

$$\begin{aligned}&\overbrace{\frac{V_s(\xi(t))}{s}-\frac{c_1}{p_1s^{2b}}\frac{2}{2n+1}r(t)^{2n+1}}^{\cdot} \leq \\ &-s\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)\left(\frac{V_s(\xi(t))}{s}-\frac{c_1}{p_1s^{2b}}\frac{2}{2n+1}r(t)^{2n+1}\right)+\\ &\frac{a_1(s)}{s^{2b}}-\frac{2c_1p_2}{s^{2b}}+\frac{2c_1}{s^{2b}}m(t)^2.\end{aligned} \tag{44}$$

We define $\rho := \frac{2}{2n+1}r^{2n+1}$, $\alpha_\circ(s)=\frac{2}{2n+1}s^{\frac{2n+1}{1+2b}}$, and

$$a_3(s) := -\frac{a_1(s)}{s^{2b}}+\frac{2c_1p_2}{s^{2b}}. \tag{45}$$

Since $a_1(s)$ goes to zero as s goes to infinity there exists $\overline{s}$ satisfying $2c_1p_2 \geq a_1(\overline{s})$. Since a_1 is also monotonic, we get $a_3(s) > 0$ for all $s > \overline{s}$. Note also that we have $\rho = \alpha_\circ(r^{1+2b})$.

By solving the differential inequality (44), we get for all $t \geq 0$,

$$\begin{aligned}\frac{V_s(\xi(t))}{s} \;\leq\; &\frac{c_1}{p_1s^{2b}}\rho(t)+\exp\left(-s\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)t\right)\\ &\left(\frac{V_s(\xi(0))}{s}-\frac{c_1}{p_1s^{2b}}\rho(0)-\frac{\frac{2c_1}{s^{2b}}\|m\|_\infty^2-a_3(s)}{s\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)}\right)\\ &+\frac{\frac{2c_1}{s^{2b}}\|m\|_\infty^2-a_3(s)}{s\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)}.\end{aligned} \tag{46}$$

By choosing

$$s > \max\left\{\frac{p_1p_2(2n+1)}{d_0}+\sqrt{\frac{c_0\hat{L}}{d_0}},\overline{s},1\right\}, \tag{47}$$

we get the existence of $\sigma_1 > 0$, solution dependent, such that for all $t \geq \sigma_1$, we have that

$$\left(\frac{V_s(\xi(0))}{s} - \frac{c_1}{p_1 s^{2b}}\rho(0) - \frac{\frac{2c_1}{s^{2b}}\|m\|_\infty^2 - a_3(s)}{s\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right)}\right) \times$$
$$\exp\left(-s\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right)t\right) \le \frac{1}{2}\,\frac{a_3(s)}{s\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right)}.$$

Then, inequality (46) becomes

$$\frac{V_s(\xi(t))}{s} \le \frac{c_1}{p_1 s^{2b}}\,\rho(t) + \frac{\frac{2c_1}{s^{2b}}\|m\|_\infty^2 - \frac{a_3(s)}{2}}{s\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right)} \qquad \forall t \ge \sigma_1 .$$

This implies, for all $t \ge \sigma_1$,

$$\rho(t) \ge \left(\frac{V_s(\xi(t))}{s} - \frac{\frac{2c_1}{s^{2b}}\|m\|_\infty^2 - \frac{a_3(s)}{2}}{s\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right)}\right)\frac{p_1 s^{2b}}{c_1}$$

and therefore, with the definition of $\alpha_\circ$, and (39)

$$\begin{aligned}\alpha_\circ\left(r(t)^{1+2b}\right) &\ge \left(\frac{V_s(\xi(t))}{s} - \frac{\frac{2c_1}{s^{2b}}\|m\|_\infty^2 - \frac{a_3(s)}{2}}{s\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right)}\right)\frac{p_1 s^{2b}}{c_1} \\ &\ge \left(\frac{V_r(\varepsilon(t))}{s\,a_2(s)} - \frac{\frac{2c_1}{s^{2b}}\|m\|_\infty^2 - \frac{a_3(s)}{2}}{s\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right)}\right)\frac{p_1 s^{2b}}{c_1}.\end{aligned} \tag{48}$$

Since $b \in (0, \frac{1}{2})$ and $r(t) \ge 1$, for all $t \ge 0$, we have that

$$\frac{d_1 p_1 p_2}{r(t)^{1+2b}} + \frac{c_0\,\hat{L}}{r^2} \le \left(d_1 p_1 p_2 + c_0\,\hat{L}\right)\frac{1}{r(t)^{1+2b}} \qquad \forall t \ge 0 .$$

Then, using Lemma 4.2 and (35), we obtain, for almost all $t \ge \sigma_1$,

$$\begin{aligned}\frac{1}{r(t)}\overbrace{V_r(\varepsilon(t))}^{\cdot} &\le -d_0 V_r(\varepsilon(t)) + \frac{\left(d_1 p_1 p_2 + c_0\,\hat{L}\right)}{r(t)^{1+2b}} V_r(\varepsilon(t)) \\ &\quad + \frac{a_1(r(t))}{r(t)^{2b}} + \frac{c_1}{r(t)^{2b}}\|m\|_\infty^2 \\ &\le -d_0 V_r(\varepsilon(t)) + \frac{\left(d_1 p_1 p_2 + c_0\,\hat{L}\right)V_r(\varepsilon(t))}{r(t)^{1+2b}} \\ &\quad + a_1(1) + c_1\|m\|_\infty^2.\end{aligned}$$

Now, for $s > \max\left\{\frac{p_1 p_2(2n+1)}{d_0} + \sqrt{\frac{c_0 \hat{L}}{d_0}}, \overline{s}, 1\right\}$, let v_0 be defined as

$$\max\left\{ s\,a_2(s)\frac{\frac{2c_1}{s^{2b}}\|m\|_\infty^2 - \frac{a_3(s)}{2}}{s\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right)} + \left(\frac{c_1}{p_1 s^{2b}}s\,a_2(s)\frac{2}{2n+1}\right) \times \left(2\frac{\left(d_1 p_1 p_2 + c_0\,\hat{L}\right)}{d_0}\right)^{\frac{2n+1}{1+2b}}, 2\frac{a_1(1) + c_1\|m\|_\infty^2}{d_0}\right\}.$$ It can be verified that we have

$$\left(\frac{v_0(s)}{s\,a_2(s)} - \frac{\frac{2c_1}{s^{2b}}\|m\|_\infty^2 - \frac{a_3(s)}{2}}{s\left(d_0 - \frac{c_0\,\hat{L}}{s^2}\right)}\right)\frac{p_1 s^{2b}}{c_1} - \alpha_\circ\left(\frac{\left(d_1 p_1 p_2 + c_0\,\hat{L}\right)\,v_0(s)}{d_0\ v_0(s) - a_1(1) - c_1\|m\|_\infty^2}\right) \ge 0$$

and since the left-hand side defines a strictly increasing function of v_0, the same holds for any $v \ge v_0(s)$. This implies $\dot{V}_r(\varepsilon(t)) < 0$ for all $t \ge \sigma_1$ such that $V_r(\varepsilon(t)) > v_0(s)$ and therefore there exists a continuous function η_1 with nonnegative values such that $\eta_1(t)$ goes to 0 as t tends to ∞ and

$$V_r(\varepsilon(t)) \le v_0(s) + \eta_1(t) \qquad \forall t \ge 0. \tag{49}$$

Boundedness of $\varepsilon(t)$ follows readily.

4.3.3 Boundedness of r

To show boundedness of $t \mapsto r(t)$, recall that (40) reads

$$\overbrace{kV_r(\varepsilon(t)) + r(t)}^{\cdot} \le k\left(d_1 p_1 p_2 + c_0\,\hat{L}\right)V_r(\varepsilon(t)) + r(t)^{1-2b}\left((k\,c_1 + 2\,p_1)\,\|m\|_\infty^2 - p_1 p_2\right) + k a_1(1) + p_1 p_2 \tag{50}$$

for almost all $t \ge 0$. Knowing that we have $\|m\|_\infty \le M_\infty$, with $k = \frac{2}{d_0 \lambda_{\min}(P)} p_1$, we choose $p_1, p_2 > 0$ in the gain adaptation law (6) to satisfy

$$M_\infty^2 \le \frac{p_1 p_2}{2(kc_1 + 2p_1)} = \frac{1}{4}\frac{1}{\frac{c_1}{d_0 \lambda_{\min}(P)} + 1} p_2 . \tag{51}$$

Then, using the inequality

$$(r_1 + r_2)^{1-2b} \le r_1^{1-2b} + r_2^{1-2b} \qquad \forall r_1,\ r_2 \ge 0$$

with $r_1 = kV_r(\varepsilon)$ and $r_2 = r$, from (50), we get, for almost all $t \ge 0$,

$$\begin{aligned}\overbrace{kV_r(\varepsilon(t)) + r(t)}^{\cdot} &\le -\frac{p_1 p_2}{2}\left(kV_r(\varepsilon(t)) + r(t)\right)^{1-2b} \\ &\quad + \frac{p_1 p_2}{2}(kV_r(\varepsilon(t)))^{1-2b} + \left(d_1 p_1 p_2 + c_0\,\hat{L}\right) kV_r(\varepsilon(t)) + k a_1(1) + p_1 p_2 .\end{aligned} \tag{52}$$

Since $t \mapsto kV_r(\varepsilon(t))$ is bounded, this inequality implies the same holds for $t \mapsto kV_r(\varepsilon(t)) + r(t)$ and $t \mapsto r(t)$.

This completes our proof of boundedness of the solution $t \mapsto (z(t), x(t), \hat{x}(t), r(t))$. Indeed, we know by assumption that $t \mapsto (z(t), x(t))$ is bounded. We have established that $t \mapsto (\varepsilon(t), r(t))$ is bounded. With the definition (28) of ε and e, we have proved the claim.

4.3.4 Results in the mean

Solving for $(e_1(t)-m(t))^2$, or equivalently, for $(\hat{y}(t)-y(t))^2$ in (6), we get, for almost all $t\geq 0$,

$$(\hat{y}(t)-y(t))^2 = \frac{\dot{r}(t)}{r(t)^{1-2b}p_1} - \frac{p_2}{r(t)^{2n+1-2b}} + p_2 .$$

It follows that

$$(\hat{y}(t)-y(t))^2 \;\leq\; p_2 \;+\; \frac{\dot{r}(t)}{r(t)^{1-2b}p_1} \;=\; p_2 \;+\; \frac{\dot{\overbrace{r(t)^{2b}}}}{2\,b\,p_1},$$

and, hence, for all $t\geq 0$ and $T>0$,

$$\frac{1}{T}\int_t^{t+T}(\hat{y}(\tau)-y(\tau))^2 d\tau \;\leq\; p_2 \;+\; \frac{1}{T}\frac{\sup_{\tau\geq t} r(\tau)^{2b}}{2\,b\,p_1}. \tag{53}$$

Inequality (13) follows from the boundedness along the solutions of r^{2b} established in Section 4.3.3.

Solving (38) between t and $t+\tau$ with

$$s > \max\left\{\sqrt{\frac{c_0\,\hat{L}}{d_0}}, 1\right\}, \tag{54}$$

gives

$$\begin{aligned}V_s(\xi(t+\tau)) \;\leq\; &\exp\left(-s\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)\tau\right)V_s(\xi(t))\\ &+ s\int_0^{\tau}\exp\left(-s\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)(\tau-\sigma)\right)\times\\ &\left(\frac{a_1(s)}{s^{2b}}+2\frac{c_1}{s^{2b}}\|m\|_\infty^2+2\frac{c_1}{s^{2b}}\left\|1-\left(\frac{\sup_{u\geq t+\sigma} r(u)}{s}\right)^n\right\|^2 (\hat{y}(t+\sigma)-y(t+\sigma))^2\right)d\sigma.\end{aligned}$$

Then, integrating between t and $t+T$, using the definition $\overline{V_s}(t)=\frac{1}{T}\int_t^{t+T}V_s(\xi(\tau))d\tau$, (53), (41), and the fact that $r(t)\geq 1$ for all t, we obtain

$$\frac{\lambda_{\min}(P)}{T}\int_t^{t+T}\xi(\tau)^\top\xi(\tau)d\tau \leq \overline{V_s}(t) \leq \frac{V_s(\xi(t))}{s\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)T} + \frac{\left(\frac{a_1(s)}{s^{2b}}+2\frac{c_1}{s^{2b}}\|m\|_\infty^2+2\frac{c_1}{s^{2b}}\sup_{\tau\geq t} r(\tau)^{2n}\,(\star_1)\right)}{\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)},$$

where $(\star_1)=p_2+\frac{1}{T}\frac{\sup_{\tau\geq t} r(\tau)^{2b}}{2\,b\,p_1}$. Then, with (32), we get

$$\frac{1}{T}\int_t^{t+T} e_i(\tau)^2 d\tau \quad\leq\quad \frac{s^{2(b+i-1)}V_s(\xi(t))}{\lambda_{\min}(P)\,s\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)T} \;+\; \frac{s^{2i}\left(a_1(s)+2c_1\|m\|_\infty^2+2c_1\sup_{\tau\geq t} r(\tau)^{2n}\,(\star_1)\right)}{\lambda_{\min}(P)\left(d_0 s^2-c_0\,\hat{L}\right)}. \tag{55}$$

To go further, we need an estimation of $\sup_{\tau\geq t} r(\tau)$. From (52), we get, with $k=\frac{2}{d_0\lambda_{\min}(P)}p_1$,

$$\begin{aligned}\sup_{\tau\geq t} r(\tau)^{1-2b} &\leq \sup_{\tau\geq t}(kV_r(\varepsilon(\tau))+r(\tau))^{1-2b}\\ &\leq \sup_{\tau\geq t}(kV_r(\varepsilon(\tau)))^{1-2b}+2+\eta_2(t)\end{aligned} \tag{56}$$

$$\left(2d_1+\frac{2c_0\,\hat{L}}{p_1p_2}\right)\sup_{\tau\geq t} kV_r(\varepsilon(\tau))+\frac{1}{p_2}\frac{4}{d_0\lambda_{\min}(P)}a_1(1),$$

where $\eta_2(t)$ tends to 0 as t goes to infinity.

Similarly, from (49), when $s>\max\left\{\frac{p_1p_2(2n+1)}{d_0}+\sqrt{\frac{c_0\hat{L}}{d_0}}, \overline{s}, 1\right\}$, we get

$$\begin{aligned}\sup_{\tau\geq t} kV_r(\varepsilon(\tau)) &\leq \frac{2}{d_0\lambda_{\min}(P)}p_1\,v_0(s) \;+\; k\,\eta_1(t)\\ &\leq \max\left\{\frac{2p_1a_2(s)\left(\frac{2c_1}{s^{2b}}\|m\|_\infty^2-\frac{a_3(s)}{2}\right)}{d_0\lambda_{\min}(P)\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)} + \left(\frac{4c_1\,s^{1-2b}\,a_2(s)}{d_0\lambda_{\min}(P)(2n+1)}\right)\left(2\frac{\left(d_1p_1p_2+c_0\,\hat{L}\right)}{d_0}\right)^{\frac{2n+1}{1+2b}},\; 2k\frac{a_1(1)+c_1\|m\|_\infty^2}{d_0}\right\} + k\,\eta_1(t).\end{aligned}$$

Using the definition of $a_3(s)$ in (45) and replacing $k=\frac{2}{d_0\lambda_{\min}(P)}p_1$, we obtain

$$\begin{aligned}\sup_{\tau\geq t} kV_r(\varepsilon(\tau)) \;\leq\; \max\Bigg\{ & p_1\frac{4a_2(s)c_1\|m\|_\infty^2+a_2(s)\,a_1(s)}{s^{2b}d_0\lambda_{\min}(P)\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)}\\ &-p_1p_2\left(\frac{2a_2(s)c_1}{s^{2b}d_0\lambda_{\min}(P)\left(d_0-\frac{c_0\,\hat{L}}{s^2}\right)}\right)\\ &+\left(\frac{4c_1\,s^{1-2b}\,a_2(s)}{d_0\lambda_{\min}(P)(2n+1)}\right)\left(2\frac{\left(d_1p_1p_2+c_0\,\hat{L}\right)}{d_0}\right)^{\frac{2n+1}{1+2b}},\\ &4p_1\frac{a_1(1)+c_1\|m\|_\infty^2}{d_0^2\lambda_{\min}(P)}\Bigg\} \;+\; k\,\eta_1(t).\end{aligned}$$

With the definitions of B_1 and B_2 in Section 2, we obtain from (55) and (56)

$$\limsup_{T\to+\infty}\limsup_{t\to+\infty}\frac{1}{T}\int_t^{t+T} e_i(\tau)^2 d\tau \;\le\; \frac{s^{2i}\left(a_1(s)+2c_1\|m\|_\infty^2\right)}{\lambda_{\min}(P)\left(d_0 s^2 - c_0\,\hat{L}\right)} + \frac{2c_1\left(B_1(s,p_1,p_2)+\frac{1}{p_2}B_2\right)^{\frac{2n}{1-2b}}}{s^{2(n-i)}\lambda_{\min}(P)\left(d_0 s^2 - c_0\,\hat{L}\right)}p_2.$$

Then, combining $s\ge 1$, (47), (51), and (54) the claim holds for

$$\begin{array}{l} s > s^* := \max\left\{\frac{p_1p_2(2n+1)}{d_0} + \sqrt{\frac{c_0\hat{L}}{d_0}}, \overline{s}, 1\right\} \\ p_2 \ge 4M_\infty^2\left(1+\frac{c_1}{d_0\lambda_{\min}(P)}\right). \end{array} \tag{57}$$

To establish Corollary 2.5, note that when $a_1(s)\equiv 0$, we have that $B_2=0$ and from (57) that

$$\limsup_{T\to+\infty}\limsup_{t\to+\infty}\frac{1}{T}\int_t^{t+T} e_i(\tau)^2 d\tau \;\le\; \frac{2s^{2i}c_1\|m\|_\infty^2}{\lambda_{\min}(P)\left(d_0 s^2 - c_0\,\hat{L}\right)} + \frac{2c_1 B_1(s,p_1,p_2)^{\frac{2n}{1-2b}}}{s^{2(n-i)}\lambda_{\min}(P)\left(d_0 s^2 - c_0\,\hat{L}\right)}p_2. \tag{58}$$

The claim follows by taking p_2 equal to its lower bound in (57).

Since for $\gamma=0$ and $m\equiv 0$ the first term of (58) vanishes, $\widetilde{B}_1$ and B_1 can be written as functions of p_1p_2 with $\widetilde{B}_1$

$$\widetilde{B}_1(s,p_1p_2) = \frac{4c_1\,s^{1-2b}\,a_2(s)}{d_0\lambda_{\min}(P)(2n+1)}\left(2\frac{\left(d_1p_1p_2+c_0\,\hat{L}\right)}{d_0}\right)^{\frac{2n+1}{1+2b}}$$

and the definition of B_1. Since, in this case, p_2 is only constrained to be positive, Corollary 2.7 readily follows.

4.4 On the case when L is known

As indicated in Remark 2.4, the first term of (58) is the bound that one would obtain when the constant vector L is known. In fact, in such a case, the analysis parameter s is not needed and r can be chosen to be a constant such that

$$d_0r^2 - c_0\,\hat{L} > 0,$$

in which case, from (37), we have, for almost all t

$$\frac{\dot{V}_r(\varepsilon(t))}{r} \le -\left(d_0 - \left(\frac{d_1p_1p_2}{r^{1+2b}} + \frac{c_0\,\hat{L}}{r^2}\right)\right)V_r(\varepsilon(t)) + \frac{a_1(r(t))}{r^{2b}} + \frac{c_1}{r^{2b}}m(t)^2.$$

Then, proceeding as to obtain (58), $\forall i\in\{1,2,\ldots,n\}$

$$\limsup_{T\to+\infty}\limsup_{t\to+\infty}\frac{1}{T}\int_t^{t+T} e_i(\tau)^2 d\tau \le \frac{r^{2i}(a_1(r)+2c_1\|m\|_\infty^2)}{\lambda_{\min}(P)\left(d_0r^2 - c_0\,\hat{L}\right)}.$$

5 Conclusion

We have shown that it is possible to design an observer to reconstruct bounded solutions of a system. We provide bounds on the mean of the error signals that can be employed to analyze performance of the observer. The main feature of the high-gain observer proposed is the on-line updated gain, which is not necessarily monotonic along solutions. This allows us, in particular, to cope with measurement noise. Even though we establish that the performance in the mean can be upper bounded as a function of the observer and analysis parameters, the price to be paid is likely a highly oscillatory behavior of the estimates. This is expected from the analysis of a closely related system studied in [18]. To improve the behavior, we have presented an adaptive procedure based on space averaging technique and involving several copies of the observer.

6 Acknowledgments

The authors would like to thank A. Ilchmann for insightful discussions on gain adaptation techniques, in particular, the results reported in [18], and to the anonymous reviewers for their helpful remarks that improved the presentation of the paper.

References

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

[2] V. Andrieu, L. Praly, and A. Astolfi. High gain observers with updated high-gain and homogeneous correction terms. *Automatica*, 45(2):422–428, 2009.

[3] A. Astolfi and L. Praly. Global complete observability and output-to-state stability imply the existence of a globally convergent observer. *Mathematics of Control, Signals, and Systems (MCSS)*, 18(1):32–65, 2006.

[4] A. A. Ball and H. K. Khalil. High-Gain observers in the presence of measurement noise: A nonlinear gain approach. In *47th IEEE Conference on Decision and Control, 2008. CDC 2008*, pages 2288–2293, 2008.

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

[6] E. Bullinger and F. Allgower. An adaptive high-gain observer for nonlinear systems. In *Proc. 36th IEEE Conference on Decision and Control*, volume 5, pages 4348–4353, 1997.

[7] C. I. Byrnes and J. C. Willems. Adaptive stabilization of multivariable linear systems. In *Proc. 23rd Conf. Decision and Control*, pages 1574–1577, 1984.

[8] J. H. Conway, R. H. Hardin, and N. J. A. Sloane. Packing lines, planes, etc.: Packings in Grassmannian spaces. *Experimental Mathematics*, 5:139–159, 1996.

[9] B. Egardt. *Stability of adaptive controllers*. Springer Verlag, 1979.

[10] J-P. Gauthier and I. Kupka. *Deterministic Observation Theory and Applications*. Cambridge University Press, 2001.

[11] J. P. Gauthier and I. A. K. Kupka. Observability and observers for nonlinear systems. *SIAM Journal on Control and Optimization*, 32(4):975–994, 1994.

[12] A. Ilchmann and D.H. Owens. Threshold switching functions in high-gain adaptive control. *IMA Journal of Mathematical Control & Information*, 8(4):409–429, 1991.

[13] P. Ioannou and P. V. Kokotovic. Instability analysis and improvement for robustness of adaptive control. *Automatica*, 20:583–594, 1984.

[14] P. Ioannou and J. Sun. *Robust adaptive control*. Prentice Hall, 1996.

[15] H. Khalil and A. Saberi. Adaptive stabilization of a class of nonlinear systems using high-gain feedback. *IEEE Transactions on Automatic Control*, 32(11):1031–1035, 1987.

[16] H. Lei, J. Wei, and W. Lin. A global observer for observable autonomous systems with bounded solution trajectories. In *44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05*, pages 1911–1916, 2005.

[17] I. M. Y. Mareels. A simple selftunning controller for stably invertible systems. *Systems and Control Letters*, 4:5–16, 1984.

[18] I. M. Y. Mareels, S. Van Gils, J. W. Polderman, and A. Ilchmann. Asymptotic dynamics in adaptive gain control. In *Advances in Control, Highlights of ECC '99*, pages 29–63, 1999.

[19] B. B. Peterson and K. Narendra. Bounded error adaptive control. *IEEE Transactions on Automatic Control*, AC-27(6):1161–1168, December 1982.

[20] L. Praly. Asymptotic stabilization via output feedback for lower triangular systems with output dependent incremental rate. *IEEE Transactions on Automatic Control*, 48(6):1103–1108, June 2003.

[21] L. Praly and Z. P. Jiang. Linear output feedback with dynamic high gain for nonlinear systems. *Systems & Control Letters*, 53(2):107–116, 2004.

[22] R. G. Sanfelice and L. Praly. A technical result for the study of high-gain observers with sign-indefinite gain adaptation. Technical report, University of Arizona, 2010. http://www.u.arizona.edu/∼sricardo/index.php?n=Main.TechnicalReports.

[23] R. G. Sanfelice and A. R. Teel. On hybrid controllers that induce input-to-state stability with respect to measurement noise. In *Proc. 44th IEEE Conference on Decision and Control and European Control Conference*, pages 4891–4896, 2005.

[24] A. Sarlette and R. Sepulchre. Consensus optimization on manifolds. *SIAM J. Control and Optimization*, 58(1):56–76, 2009.

[25] L. K. Vasiljevic and H. K. Khalil. Differentiation with High-Gain observers the presence of measurement noise. In *45th IEEE Conference on Decision and Control*, pages 4717–4722, 2006.